(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/04645 A2

(51) International Patent Classification[7]: C12N 15/56, 9/24, C07K 16/40, C12Q 1/68, G01N 33/68

(21) International Application Number: PCT/EP01/08094

(22) International Filing Date: 12 July 2001 (12.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00202442.0 12 July 2000 (12.07.2000) EP

(71) Applicant *(for all designated States except US)*: **VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW** [BE/BE]; Rijvisschestraat 120, B-9052 Zwijnaarde (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DAVID, Guido** [BE/BE]; Koningin Elisabethlaan 11, B-Belgium Leuven (BE). **DÜRR, Joachim** [DE/BE]; Bergstraat 19, B-1970 Wezembeek-Oppem (BE).

(74) Common Representative: **VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW**; Rijvisschestraat 120, B-9052 Zwijnaarde (BE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/04645 A2

(54) Title: A SECOND HUMAN HEPARANASE, AND SPLICE VARIANTS THEREOF, WITH A PREDOMINANT EXPRESSION IN SKELETAL MUSCLE, HEART AND PANCREAS

(57) Abstract: The present invention relates to the field of carbohydrates and more specifically to the field of heparan sulphate proteoglycans. A novel polynucleotide has been identified encoding heparanase activity. This is the second gene encoding heparanase catalytic activity which is identified in humans. Several splice variants of said gene have been identified with a specific expression pattern in skeletal muscle, heart and pancreas.

A second human heparanase, and splice variants thereof, with a predominant expression in skeletal muscle, heart and pancreas.

Field of the invention

The present invention relates to the field of carbohydrates and more specifically to the field of heparan sulphate proteoglycans. A novel polynucleotide has been identified encoding heparanase activity. This is the second gene encoding heparanase catalytic activity which is identified in humans. Several splice variants of said gene have been identified with a specific expression pattern in skeletal muscle, heart and pancreas.

Background of the invention

Heparan sulfate proteoglycans (HSPGs) are ubiquitous macromolecules associated with the cell surface and extracellular matrix (ECM) of a wide range of cells of vertebrate and invertebrate tissues (David G., (1993), *Faseb J.* 7, 1023).

The basic HSPG structure consists of a protein core to which several linear heparan sulfate (HS) chains are covalently attached. The HS chains are typically composed of repeating hexuronic acid and D-glucosamine disaccharide units that are substituted to a varying extent with N-and O-linked sulfate moieties and N-linked acetyl groups (Kjellen L. & Lindahl U., (1991), *Annu. Rev. Biochem.* 60, 443). The ability of HS to interact with ECM macromolecules such as collagen, laminin and fibronectin, and with different attachment sites on plasma membranes indicates this proteoglycan is essential in the self-assembly and insolubility of ECM components, as well as in cell adhesion and locomotion (Rapraeger A.C., (1993), *Curr. Opin. Cell Biol.* 5, 844).

In addition, biochemical studies and cell-culture assays have implicated heparan sulphate proteoglycans (HSPGs) as co-receptors (Bernfield M. *et al.*, (1992), *Annu. Rev. Cell Biol.* 8, 365) in processes ranging from mechanical support to functions in adhesion, motility, proliferation, differentiation and morphogenesis. Enormous structural heterogeneity can be generated through specific heparin sulphate (HS) chain modifications during their biosynthesis, as well as from the diverse nature of their core proteins. Furthermore, biochemical studies have shown how interactions of cell-surface HSPGs can be specific at various levels. Tissue-specific isoforms of the polymerizing and chain-modifying enzymes can produce HS chains with distinct sequences and macroscopic organizations (Habuchi H., Habuchi O. & Kimata K., (1998), *Trends Glycosci. Glycotechnol.* 10, 65). Various ligands and their receptors

can show selectivity in their binding affinities to distinct HS chain structures (Lyon M. & Gallagher J.T., (1998), *Matrix Biol.* 17, 485). These chains can be attached to one or more specific core proteins, each with a distinctive tissue-specific expression pattern and cellular localization. Also, HSPGs can be selectively shed from the cell surface to yield soluble effectors.

HSPGs are prominent components of blood vessels (Wight T.N., (1989), *Arteriosclerosis* 9, 1). In capillaries they are found mainly in the subendothelial basement membrane, where they support the vascular endothelium and stabilize the structure of the capillary wall. Cleavage of HS therefore plays a decisive part in the extravasation of blood-borne cells. In fact, expression of HS-degrading endoglycosidases, commonly called 'heparanases', correlates with the metastatic potential of mouse lymphoma, fibrosarcoma and melanoma cell lines (Vlodavsky I. et al. (1983), *Cancer Res.* 43, 2704). Moreover, increased levels of heparanase were detected in sera (Nakajima et al. (1983), *Science* 220, 611) and urine of metastatic tumor-bearing animals and cancer patients. Treatment of experimental animals with heparanase inhibitors (for example, non-anticoagulant species of low-molecular-weight heparin and polysulfated saccharides) considerably reduced the incidence of lung metastases by melanoma, Lewis lung carcinoma and mammary adenocarcinoma cells (Vlodavsky I., *et al.*, (1995), *Invasion Metastasis* 14, 290). Heparanase-inhibiting molecules also inhibit T cell-mediated delayed-type hypersensitivity and experimental autoimmune encephalomyelitis and adjuvant arthritis (Vlodavsky I., *et al.*, (1992), *Invasion Metastasis* 12, 112), reflecting a role in cell diapedesis and extravasation associated with inflammation and autoimmunity. Endoglycosidases, mainly endo-beta-D-glucuronidases, capable of partially depolymerising HS chains, have been demonstrated in a variety of cells and tissues. In addition to being involved in the remodelling of the ECM and in the egress of cells from the vasculature, heparanases are implicated in angiogenesis, tissue repair, inflammation, diabetes, asthma and lipid metabolism by releasing HS-bound growth factors and enzymes such as for example basic fibroblast growth factor (bFGF) and lipoprotein lipase. Research on the involvement of heparanase in these processes has been handicapped by the lack of purified enzyme preparations and appropriate molecular probes and antibodies to explore a causative role for heparanase in normal and pathological processes.

The cDNA sequences of the first mammalian heparanase (heparanase-1 or hep1) from human placenta and platelets have recently been reported (Vlodavsky I. *et al.*, (1999),

Nat. Med. 5, 793; Hulett *et al.*, (1999), *Nat. Med.* 5, 803; Kussie *et al.*, (1999), *Biochem. Biophys. Res. Commun.* 261, 183), US 5968822 and PCT/EP99/00777). Since it is generally accepted that all cell types express one or multiple or even specific heparanases it is possible that many more structurally distinct genes exist that encode other heparanases. It is an object of the present invention to provide a second human heparanase cDNA (hep2) and several splice variants thereof. The novel heparanase cDNA has a specific expression level that is mainly restricted to the heart, pancreas and skeletal muscle. Furthermore, contrary to what is reported in WO 01/21814, we have shown that the isolated hep2 is catalytically active.

Figure and table legends

Fig. 1: Splice variants of heparanase 2.

Four different splice variants of heparanase 2 could be identified by RACE experiments or touchdown PCR (see table 1). Touchdown PCR with primers PR 23-50/23-41 (Table 3) flanking the splice variant sites A and B was performed to verify the presence of these different splice variants in the libraries and in RT-mRNA of CaCo2 cells.

Table 1: Presence of the various heparanase 2 splice variants in different libraries and CaCo2-cells.

The four different splice variants as drawn in Fig.1 were identified via RACE experiments and verified by PCR using primer pairs flanking the splice variant site AB. Note that the longest transcript Hep2AB could only be identified via RACE. The calculated molecular weight (Mw) according to the primary amino acid sequence is listed.

Table 2: Primer combinations used in RACE experiments. Only experiments that gave new sequence information are listed.

Table 3: Sequence of the Primers used

Table 4: Expression patterns of heparanase 1 and heparanase 2

Multiple tissue northerns (clontech) were probed with the C-terminal portion of heparanase 2 amplified with primers PR22-10/PR22-11.

Hep2 is also expressed in the human colon carcinoma cell line CaCo2 as shown by RT-PCR.

Aims of the invention

The current invention aims at providing a polynucleotide and functional fragments thereof, designated as Hep2, encoding a polypeptide having heparanase catalytic activity, vectors including the same, transduced cells expressing Hep2 and a recombinant polypeptide or functional fragment thereof having heparanase 2 activity. In addition the invention aims at providing specific splice variants of Hep2. The invention further aims at providing a screening method for molecules that have a potential at antagonizing or agonizing the heparanase 2 catalytic activity. The invention also aims at providing molecules obtained from the screening assay that can be used for the manufacture of a medicament. Another aim of the invention is to provide an antibody that specifically recognizes and binds to a Hep2 polypeptide. Finally, the invention aims at using polymorphisms of the Hep2 sequence to identify individuals having a predisposition to acquire diseases resulting from a shortage or excessive activity of heparanase 2 activity.

Detailed description of the invention

According to the present invention there is provided a polynucleotide, referred to hereinbelow as Hep2, Hep2 cDNA or Hep2 gene encoding a polypeptide having heparanase catalytic activity, vectors including the same, transduced host cells expressing said heparanase and a recombinant protein having said heparanase catalytic activity.

A number of terms and expressions are used throughout the detailed description and, to facilitate the understanding thereof, the following definitions are provided:

As used herein, the words "polynucleotide" may be interpreted to mean the DNA and cDNA sequence as detailed by Yoshikai et al. (1990) *Gene* 87:257, with or without a promoter DNA sequence as described by Salbaum et al. (1988) *EMBO J.* 7(9):2807.

As used herein, "fragment" refers to a polypeptide or polynucleotide of at least about 9 amino acids or 27 base pairs, typically 50 to 75, or more amino acids or base pairs, wherein the polypeptide contains an amino acid core sequence. A fragment may be for example a truncated Hep2 isoform, modified Hep2 isoform (as by amino acid substitutions, deletions, or additions outside of the core sequence), or other variant polypeptide sequence, but is not a naturally-occurring Hep2 isoform that is present in a human individual. If desired, the fragment may be fused at either terminus to additional

amino acids or base pairs, which may number from 1 to 20, typically 50 to 100, but up to 250 to 500 or more. A "functional fragment" means a polypeptide fragment possessing the biological property of having heparanase catalytic activity or a polynucleotide fragment encoding heparanase catalytic activity.

We believe that we are the first to show catalytic activity of hep2. Furthermore, the optimal pH-range for catalytically active hep2 is different from hep1, and hep2 cleaves HS chains in a different manner than hep1. WO 01/21814 describes the isolation of hep2 but no catalytic activity is demonstrated. Also in WO 01/21814 the amino acid op hep2 has an extension of 10 amino acids as compared to SEQ ID NO: 2 which depicts the amino acids of hep2 of the present invention. There are several reasons why we believe that SEQ ID NO: 2 is the correct amino acid sequence of hep2 and that amino terminal extensions of SEQ ID NO: 2 lead to catalytically inactive hep2: (1) *in silico* analysis of SEQ ID NO: 1 shows that the sequence context upstream of the first ATG-initiation codon figures a pyrimidine (T) at the minus-3 position, which makes it a far less than optimal Kozak sequence. However, the second ATG-initiation codon, in contrast, is preceded by a GCC codon, figuring a purine (G) at the minus-3 position and also perfectly matching the Kozak consensus sequence (G/ACCATG) at the minus-2 and minus-1 positions, (2) translation of the protein from the first ATG codon on would result in a protein that lacks a typical signal sequence. Translation starting at the second ATG, in contrast, results in a protein with a typical signal sequence, with a signal peptidase cleavage site located between amino acids Ser^{28} and Ser^{29} (based on the psort program http://psort.nibb.ac.jp/) and (3) heparanase 2 is detected in the culture medium of transfected cells and we believe that the amino acid sequence is as depicted in SEQ ID NO: 2 and (4) SEQ ID NO: 2 and its splice variants described herein clearly possess catalytic activity.

According to further features in preferred embodiments of the invention described below, provided is a polynucleotide sequence which includes polynucleotide fragments encoding polypeptides having heparanase catalytic activity.

According to still further features in the described preferred embodiments the polynucleotide fragment includes nucleotides 85-1833 of SEQ ID NO: 1, which encodes the entire human heparanase 2 enzyme (Hep2).

According to still further features four different Hep2 cDNAs (splice variants), generated as a result of alternative splicing, encoding human Hep 2AB (set forth in SEQ ID NO: 2), Hep 2A (set forth in SEQ ID NO: 4), Hep 2B (set forth in SEQ ID NO:

6) and Hep 2- - (set forth in SEQ ID NO: 8) were isolated containing 582, 524, 528 and 470 amino acid residues, respectively.

According to still further features in the described preferred embodiments the polynucleotide sequence which encodes the polypeptide having heparanase 2 activity shares at least 60% homology, preferably at least 70% homology more preferably at least 80% homology, most preferably at least 90% homology with SEQ ID NO:1.

Homology is determined using default parameters of a DNA sequence analysis software package developed by the Genetic Computer Group (GCG) at the University of Wisconsin.

According to still further features in the described preferred embodiments the polynucleotide fragment according to the present invention includes a portion (fragment) of SEQ ID NO:1 which encodes a polypeptide having the heparanase catalytic activity.

According to still further features in the described preferred embodiments the polypeptide encoded by the polynucleotide fragment includes an amino acid sequence as set forth in SEQ ID NO:2 or a functional fragment thereof.

According to still further features in the described preferred embodiments the polynucleotide sequence encodes a polypeptide having heparanase activity, which shares at least 60% homology, preferably at least 70% homology, more preferably at least 80% homology, most preferably at least 90% homology with SEQ ID NO: 2.

According to still further features in the described preferred embodiments the polynucleotide fragment encodes a polypeptide having heparanase activity, which may therefore be allelic, species and/or induced variant of the amino acid sequence set forth in SEQ ID NO:2. It is understood that any such variant may also be considered a homolog. The term "allelic variant" is used herein to denote any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequence. The term allelic variant is also used herein to denote a protein encoded by an allelic variant of a gene.

According to still further features in the described preferred embodiments provided is a single stranded polynucleotide fragment which includes a polynucleotide sequence complementary to at least a portion of a polynucleotide strand encoding a polypeptide having heparanase 2 catalytic activity as described above.

The expression pattern of Hep2 mRNA in various tissues and cell lines was investigated using RT-PCR. Hep2 was found to be expressed in tissues and cells previously known to have heparanase activity. However, the observed expression pattern is completely different from the previously identified heparanase (Hulett *et al*, (1999), *Nat. Med.* 5, 803). Regarding the organ specific expression of Hep2 which is mainly observed in skeletal muscle, pancreas and heart and to a lesser extent in kidney and lung, it is expected that excessive Hep2, or a specific splice variant thereof, or a shortage of Hep2, or a specific splice variant thereof, is implicated in specific disease processes of said organs. For example in the lung, Hep2, or a specific splice variant thereof, can be a target for the treatment of asthmatic patients. In said patients it is observed that an enhanced proteoglycan deposition and/or turnover contributes to the airway wall remodelling observed in asthmatics (Huang et al. (1999) *Am J Respir Crit Care Med*, 160, 725 and Roberts and Burke (1998) *Can Respir J* 5, 48). In the pancreas, Hep2, or a specific splice variant thereof, can be a target for the treatment of human pancreatic cancer. It has recently been demonstrated that the heparan sulfate proteoglycan, glypican-1, is strongly overexpressed in human pancreatic cancer (Kleeff et al. (1998) *J Clin Invest*, 102, 1662) and hence it is proposed that glypican-1 plays an essential role in the responses of pancreatic cancer cells to certain mitogenic stimuli. Consequently, use of a medical preparation comprising an agonist of Hep2, or a specific splice variant thereof, could be useful for the treatment of pancreatic cancer. Also skeletal muscle fibers are surrounded by an extracellular matrix which is composed of glycoproteins, collagen, and proteoglycans. It has been observed that an upregulation of proteoglycans is positively correlated with skeletal muscle regeneration (Caceres et al. (2000) *Eur J Cell Biol* 79, 173). Also here it can be envisaged that the inhibition of Hep2, or a specific splice variant thereof, can be useful for the treatment of dystrophic muscular diseases.

In certain heart diseases such as myxomatous mitral valve defects it is observed that proteoglycans are abnormally high around degenerated elastic fibres and collagen fibres (Akhtar el al. (1999) *Cardiovasc Pathol* 4, 191). Also in experimental induced myocardial infarction models an increase of specific proteoglycans is observed (Doi et al ((2000) Pathol Res Pract 196, 23). It can be expected that a medical preparation comprising an agonist of Hep2 can be useful for the prophylaxis and/or treatment of heart diseases.

In what follows below, a short description of disease processes, that can occur because of excessive or a shortage of heparanase 2, or a specific splice variant thereof, catalytic activity, is given for the sake of clarity and completeness.

a) Possible involvement of heparanase2, or a specific splice variant thereof, in tumor cell invasion and metastasis:

Circulating tumor cells arrested in the capillary beds of different organs must invade the endothelial cell lining and degrade its underlying basement membrane (BM) in order to invade into the extravascular tissue(s) where they establish metastasis (Nicolson G.L., (1988), *Cancer Met. Rev.* 7, 143). Once located between endothelial cells and the BM, the invading cells must degrade the subendothelial glycoproteins and proteoglycans of the BM in order to migrate out of the vascular compartment. Several cellular enzymes (e.g., collagenase IV, plasminogen activator, cathepsin B, elastase, etc.) are thought to be involved in degradation of BM (Liotta L.A. et al (1983), *Lab. Invest.* 49, 639). Among these enzymes is a beta-D-glucuronidase (heparanase) that cleaves HS at specific intrachain sites (Vlodavsky I. *et al.*, (1992), *Invasion & Metastasis* 12, 112). Expression of a HS degrading heparanase was found to correlate with the metastatic potential of mouse lymphoma (Vlodavsky I. *et al.*, (1983), *Cancer Res.* 43, 2704), fibrosarcoma and melanoma (Nakajima M. *et al*, (1988), *J. Cell. Biochem.* 36, 157) cells. Moreover, elevated levels of heparanase were detected in sera from metastatic tumor bearing animals and melanoma patients (Nakajima M. *et al*, (1988), *J. Cell. Biochem.* 36, 157) and in tumor biopsies of cancer patients (Vlodavsky I. *et al*, (1988), *J. Med.* 24, 464).

b) Possible involvement of heparanase, or a specific splice variant thereof, in tumor angiogenesis:

Fibroblast growth factors are a family of structurally related polypeptides characterized by high affinity to heparin (Burgess W.H. & Maciag T., (1989), *Annu. Rev. Biochem.* 58, 575). They are highly mitogenic for vascular endothelial cells and are among the most potent inducers of neovascularization (Burgess W.H. & Maciag T., (1989), *Annu. Rev. Biochein.* 58, 575). Basic fibroblast growth factor (bFGF) has been extracted from the subendothelial ECM produced in vitro (Vlodavsky I. *et al*, (1987), *Proc.Natl.Acad.Sci.USA* 84, 2292) and from basement membranes of the cornea (Folkman J. *et al,* (1980), *Am. J. Pathol.* 130, 393). suggesting that ECM may serve as a reservoir for bFGF. Immunohistochemical staining revealed the localization of bFGF in basement membranes of diverse tissues and blood vessels (Cardon-Cardo C. *et al*,

(1990), *Lab Invest.* 63, 832). Despite the ubiquitous presence of bFGF in normal tissues, endothelial cell proliferation in these tissues is usually very low, suggesting that bFGF is somehow sequestered from its site of action. Studies on the interaction of bFGF with ECM revealed that bFGF binds to HSPG in the ECM and can be released in an active form by HS degrading enzymes (Ishai-Michaeli R. *et al*, (1992), *Biochemistry* 31, 2080). It was demonstrated that heparanase activity expressed by platelets, mast cells, neutrophils and lymphoma cells is involved in release of active bFGF from ECM and basement membranes (Ishai-Michaeli R. *et al*, (1990), *Cell Reg.* 1, 833), suggesting that heparanase activity may not only function in cell migration and invasion, but may also elicit an indirect neovascular response. These results suggest that the ECM HSPG provides a natural storage depot for bFGF and possibly other heparin-binding growth promoting factors (Vlodavsky I. *et al*, (1991), *Trends Biochem. Sci* 16, 268). Displacement of bFGF from its storage within basement membranes and ECM may therefore provide a novel mechanism for induction of neovascularization in normal and pathological situations. Recent studies indicate that heparin and HS are involved in binding of bFGF to high affinity cell surface receptors and in bFGF cell signaling (Spivak-Kroizman T. *et al*, (1994), *Cell* 79, 1015). Moreover, the size of HS required for optimal effect was similar to that of HS fragments released by heparanase (Ornitz D. *et al* (1995), *Science* 268, 432). Similar results were obtained with vascular endothelial cells growth factor (VEGF) (Gitay-Goren H. *et al*, (1992), *J. Biol. Chem.* 267, 6093), suggesting the operation of a dual receptor mechanism involving HS in cell interaction with heparin-binding growth factors. It is therefore proposed that restriction of endothelial cell growth factors in ECM prevents their systemic action on the vascular endothelium, thus maintaining a very low rate of endothelial cells turnover and vessel growth. On the other hand, release of bFGF from storage in ECM as a complex with HS fragment, may elicit localized endothelial cell proliferation and neovascularization in processes such as wound healing, inflammation and tumor development (Vlodavsky I. *et al*, (1991), *Trends Biochem. Sci* 16, 268).

c) Possible involvement of heparanase, or a specific splice variant thereof, in the immune system:

Heparanase activity correlates with the ability of activated cells of the immune system to leave the circulation and elicit both inflammatory and autoimmune responses. Interaction of platelets, granulocytes, T and B lymphocytes, macrophages and mast cells with the subendothelial ECM is associated with degradation of HS by a specific

heparanase activity (Vlodavsky I. *et al.*, (1992), *Invasion & Metastasis* 12, 112). The enzyme is released from intracellular compartments (e.g., lysosomes, specific granules. etc.) in response to various activation signals (e.g., thrombin, calcium ionophore, immune complexes. antigens. mitogens. etc.). suggesting its regulated involvement in inflammation and cellular immunity.

d) Involvement of heparanase, or a specific splice variant thereof, in other potential disease processes:

Apart from its involvement in tumor cell metastasis, inflammation and autoimmunity, mammalian heparanase may be applied to modulate: bioavailability of heparin-binding growth factors (Ruoslahti E. and Yamaguchi Y., (1991), *Cell* 64, 867); cellular responses to heparin-binding growth factors (e.g., bFGF, VEGF) and cytokines (IL-8) (Gitay-Goren H. *et al*, (1992), *J. Biol. Chem.* 267, 6093); cell interaction with plasma lipoproteins (Eisenberg S. *et al*, (1992), *J. Clin. Invest.* 90, 2013); cellular susceptibility to certain viral and some bacterial and protozoa infections (Putnak J.R. *et al*, (1997) *Nature Medicine* 3, 828); and disintegration of amyloid plaques (Narindrasorasak S. *et al*, (1991), *J. Biol. Chem.* 266, 12878). Heparanase-2, or a specific splice variant thereof, may thus prove useful for conditions such as wound healing, angiogenesis, restenosis, atherosclerosis, inflammation, neurodegenerative diseases and viral infections. Mammalian heparanase-2, or a specific splice variant thereof, can be used to neutralize plasma heparin as a potential replacement of protamine. Anti-heparanase-2, or a specific splice variant thereof, antibodies may be applied for immunodetection and diagnosis of micrometastases, autoimmune lesions and renal failure in biopsy specimens, plasma samples, and body fluids. Common use in basic research is expected.

The presence of heparan sulphate on cell surfaces has been shown to be the principal requirement for the binding of Herpes Simplex (Shieh M.T. *et al*, (1992), *J. Cell Biol.* 116, 1273) and Dengue (Chen Y. *et al*, (1997), *Nature Medicine* 3, 866) viruses to cells and for subsequent infection of the cells. Removal of the cell surface heparan sulfate by heparanase-2, or a specific splice variant thereof, may therefore abolish virus infection. In fact, treatment of cells with bacterial heparitinase (degrading heparan sulphate) or heparinase (degrading heparin/heparan sulphate) reduced the binding of two related animal herpes viruses to cells and rendered the cells at least partially resistant to virus infection (Shieh M.T. *et al*, (1992), *J. Cell Biol.* 116, 1273). There are

some indications that the cell surface heparan sulphate is also involved in HIV infection (Putnak J.R. *et al*, (1997) *Nature Medicine* 3, 828).

Heparan sulfate proteoglycans were identified in the prion protein amyloid plaques of Gerstmann-Straussler Syndrome, Creutzfeldt-Jakob disease and Scrapie (Narindrasorasak S. *et al*, (1991), *J. Biol. Chem.* 266, 12878). Heparanase-2, or a specific splice variant thereof, may disintegrate these amyloid plaques which are also thought to play a role in the pathogenesis of Alzheimer's disease.

Proliferation of arterial smooth muscle cells (SMCs) in response to endothelial injury and accumulation of cholesterol rich lipoproteins are basic events in the pathogenesis of atherosclerosis and restenosis (Ross R., (1993), *Nature* 362, 801). Apart from its involvement in SMC proliferation (i.e., low affinity receptors for heparin-binding growth factors), HS is also involved in lipoprotein binding, retention and uptake (Zhong-Sheng J. *et al*, (1993), *J. Biol. Chem.* 268, 10160). It was demonstrated that HSPG and lipoprotein lipase participate in a novel catabolic pathway that may allow substantial cellular and interstitial accumulation of cholesterol rich lipoproteins (Eisenberg S. *et al*, (1992), *J. Clin. Invest.* 90, 2013). The latter pathway is expected to be highly atherogenic by promoting accumulation of apoB and apoE rich lipoproteins (i.e. LDL, VLDL, chylomicrons), independent of feed back inhibition by the cellular sterol content. Removal of SMC HS by heparanase-2, or a specific splice variant thereof, is therefore expected to inhibit both SMC proliferation and lipid accumulation and thus may halt the progression of restenosis and atherosclerosis.

According to still further features in the described preferred embodiments provided is a vector including a polynucleotide sequence encoding a polypeptide having heparanase-2 catalytic activity.

The vector may be of any suitable type including, but not limited to, a phage, virus, plasmid, phagemid, cosmid, bacmid or even an artificial chromosome. The polynucleotide sequence encoding a polypeptide having heparanase 2 catalytic activity may include any of the above described polynucleotide fragments.

As used herein, the term "regulatory element" refers to a genetic element which controls some aspect of the expression of nucleic acid sequences. For example, a promoter is a regulatory element that facilitates the initiation of transcription of an operably linked coding region. Other regulatory elements are splicing signals, polyadenylation signals, termination signals, etc.) Transcriptional control signals in eucaryotes comprise "promoter" and "enhancer" elements. Promoters and enhancers

consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription [Maniatis, T. et al., *Science* 236:1237 (1987)]. Promoter and enhancer elements have been isolated from a variety of eukaryotic sources including genes in yeast, insect and mammalian cells and viruses (analogous control elements, i.e., promoters, are also found in procaryotes). The selection of a particular promoter and enhancer depends on what cell type is to be used to express the protein of interest. Some eukaryotic promoters and enhancers have a broad host range while others are functional in a limited subset of cell types [for review see Voss, S. D. et al., Trends Biochem. Sci., 11:287 (1986) and Maniatis, T. et al., supra (1987)].

The term "recombinant DNA vector" as used herein refers to DNA sequences containing a desired coding sequence and appropriate DNA sequences necessary for the expression of the operably linked coding sequence in a particular host organism (e.g. mammal). DNA sequences necessary for expression in procaryotes include a promoter, optionally an operator sequence, a ribosome binding site and possibly other sequences. Eukaryotic cells are known to utilize promoters, polyadenlyation signals and enhancers.

According to still further features in the described preferred embodiments provided is a host cell that includes an exogenous polynucleotide fragment including a polynucleotide sequence encoding a polypeptide having heparanase 2 catalytic activity.

The exogenous polynucleotide fragment may be any of the above described fragments. The host cell may be of any type such as prokaryotic cell, eukaryotic cell, a cell line, or a cell as a portion of a multicellular organism (e.g., cells of a transgenic organism).

According to still further features in the described preferred embodiments provided is a recombinant protein including a polypeptide having heparanase 2 catalytic activity. The recombinant protein may be purified by any conventional protein purification procedure close to homogeneity and/or be mixed with additives. The recombinant protein may be manufactured using recombinant expression systems comprising bacterial cells, yeast cells, animal cells, insect cells, plant cells or transgenic animals or plants.

According to still further features in the described preferred embodiments provided is a pharmaceutical composition comprising as an active ingredient a recombinant protein having heparanase 2 catalytic activity.

According to still further features in the described preferred embodiments provided is a medical equipment comprising a medical device containing, as an active ingredient a recombinant protein having heparanase 2 catalytic activity.

According to still further features in the described preferred embodiments provided is a heparanase overexpression system comprising a cell overexpressing heparanase 2 catalytic activity. The cell may be a host cell transiently or stably transfected or transformed with any suitable vector which includes a polynucleotide sequence encoding a polypeptide having heparanase activity and a suitable promoter and enhancer sequences to direct expression of heparanase 2. However, the overexpressing cell may also be a product of an insertion (e.g. via homologous recombination) of a promoter and/or enhancer sequence downstream to the endogenous heparanase gene of the expressing cell, which will direct overexpression from the endogenous gene. The term 'overexpression' as used herein refers to a level of expression which is higher than a basal level of expression typically characterizing a given cell under otherwise identical conditions.

The present invention can be used to develop assays to identify molecules to inhibit for example tumor cell metastasis, inflammation and autoimmunity. The identification of the Hep2 gene encoding for heparanase 2 enzyme enables the production of a recombinant enzyme in heterologous expression systems, alternatively the Hep 2 can be purified from cell lines expressing naturally Hep 2 by methods known in the art.

Heparanase inhibitors, mainly based on heparin and similar polysaccharides, in the prior art have been shown to inhibit tumour growth and/or metastasis, angiogenesis and vascular damage in some cases in experimental models. The availability of large quantities of recombinant enzyme and sensitive functional assays will facilitate the design and testing of better and more selective inhibitors. However, as with many enzyme systems, balance is essential, and the finding that for example bFGF signalling is facilitated when bound to cell surface HS cautions that use of heparanase inhibitors may shift the balance from free bFGF to HS-bFGF, alter recycling and degradation pathways and enhance rather than inhibit cellular activation. Careful evaluation of possible adverse effects on normal physiological functions will also be imperative. Alternatively, agonists can be identified with the screening assay and can be used to manufacture a medicament, or in gene therapy, for treatment of diseases where a shortage of Hep 2 occurs.

Therefore the invention also provides methods for identifying compounds or molecules which bind on the Hep2 polypeptide or a functional fragment thereof and interfere with its catalytic heparanase2 activity. These methods are also referred to as 'drug screening assays' or 'bioassays' and typically include the step of screening a candidate/test compound or agent for the ability to interact with Hep2. Candidate compounds or agents, which have this ability, can be used as drugs to combat or prevent for example tumour invasion.

Small molecules, e.g. small organic molecules, and other drug candidates can be obtained, for example, from combinatorial and natural product libraries. Random peptide libraries consisting of all possible combinations of amino acids attached to a solid phase support may also be used to identify peptides that are able to bind to specific ligands (Lam KS et al., 1991, Nature 354, 82). Identification of molecules that are able to bind to Hep2 may be accomplished by screening a peptide library with recombinant soluble Hep2 protein.

One possibility of assays are cell-free assays, which include the steps of combining Hep2 and a candidate/test compound, e.g., under conditions which allow for interaction of (e.g. binding of) the candidate/test compound with Hep2 to form a complex, and detecting the formation of a complex, in which the ability of the candidate compound to interact with Hep2 is indicated by the presence of the candidate compound in the complex. Formation of complexes between the Hep2 and the candidate compound can be quantitated, for example, using standard immunoassays. The Hep2 employed in such a test may be free in solution or affixed to a solid support. Alternatively, cell-based assays may be used to identify compounds that can interact with Hep2, an example but not limited to this, is yeast-two-hybrid and its derivatives or the specific infection by phages of cells, expressing Hep2, by phages expressing hybrid molecules that can bind to Hep2.

To perform the above described drug screening assays, it is feasible to immobilize Hep2 or its (their) target molecule(s) to facilitate separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Interaction (e.g., binding of) of Hep2 to a target molecule, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and microcentrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, Hep2 tagged can be adsorbed onto Ni-

NTA microtiter plates, or Hep2-ProtA fusions adsorbed to IgG, which are then combined with the cell lysates (e.g., ^{35}S-labeled) and the candidate compound, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the plates are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly, or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-PAGE, and the level of Hep2 binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques. Other techniques for immobilizing protein on matrices can also be used in the drug screening assays of the invention. For example, either Hep2 or its target molecules can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated Hep2 can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive to Hep2 but which do not interfere with binding of the protein to its target molecule can be derivatized to the wells of the plate, and Hep2 trapped in the wells by antibody conjugation. As described above, preparations of a Hep2-binding protein and a candidate compound are incubated in the Hep2-presenting wells of the plate, and the amount of complex trapped in the well can be quantitated. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive to the Hep2-target molecule, or which are reactive to Hep2 and compete with the target molecule; as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the target molecule. Another technique for drug screening which provides for high throughput screening of compounds having suitable binding affinity to Hep2 is described in detail in "Determination of Amino Acid Sequence Antigenicity" by Geysen HN, WO 84/03564, published on 13/09/84. In summary, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The protein test compounds are reacted with fragments of Hep2 and washed. Bound Hep2 is then detected by methods well known in the art. Purified Hep2 can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support. This invention also

contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding Hep2 specifically compete with a test compound for binding Hep2. In this manner, the antibodies can be used to detect the presence of any protein, which shares one or more antigenic determinants with Hep2.

In patent WO 00/03036 qualitative and quantitative methods of testing molecules for its potential at inhibiting heparanase in the presence of a heparanase substrate are fully described.

The present invention also relates to molecules that can be used to neutralize the activity of Hep2 by interfering with its synthesis and/or translation. By molecules it is meant peptides, proteins, organic molecules and carbohydrates. More specifically, the invention is directed to antagonists of Hep2 such as anti-Hep2 antibodies and functional derivatives derived thereof, anti-sense RNA and DNA molecules and ribozymes that function to inhibit the translation of Hep2.

By 'synthesis' it is meant trancription of Hep2. Small molecules can bind on the promoter region of Hep2 and inhibit binding of a transcription factor or said molecules can bind said transcription factor and inhibit binding to the Hep2-promoter. By Hep2 it is meant also its isoforms, which occur as a result of alternative splicing, and allelic variants thereof. As a result of alternative splicing, four Hep2 RNAs encoding human Hep 2AB (set forth in SEQ ID NO:2), Hep 2A (set forth in SEQ ID NO: 4), Hep 2B (set forth in SEQ ID NO: 6) and Hep 2- - (set forth in SEQ ID NO: 8) isoform precursors containing 582, 524, 528 and 470 amino acid residues, respectively, have been identified.

The term 'antibody' or 'antibodies' relates to an antibody characterized as being specifically directed against Hep2 or any functional derivative thereof, with said antibodies being preferably monoclonal antibodies; or an antigen-binding fragment thereof, of the $F(ab')_2$, F(ab) or single chain Fv type, or any type of recombinant antibody derived thereof. These antibodies of the invention, including specific polyclonal antisera prepared against Hep2 or any functional derivative thereof, have no cross-reactivity to others proteins. The monoclonal antibodies of the invention can for instance be produced by any hybridoma liable to be formed according to classical methods from splenic cells of an animal, particularly of a mouse or rat immunized against Hep2 or any functional derivative thereof, and of cells of a myeloma cell line, and to be selected by the ability of the hybridoma to produce the monoclonal antibodies recognizing Hep2 or any functional derivative thereof which have been

initially used for the immunization of the animals. The monoclonal antibodies according to this embodiment of the invention may be humanized versions of the mouse monoclonal antibodies made by means of recombinant DNA technology, departing from the mouse and/or human genomic DNA sequences coding for H and L chains or from cDNA clones coding for H and L chains. Alternatively the monoclonal antibodies according to this embodiment of the invention may be human monoclonal antibodies. Such human monoclonal antibodies are prepared, for instance, by means of human peripheral blood lymphocytes (PBL) repopulation of severe combined immune deficiency (SCID) mice as described in PCT/EP 99/03605 or by using transgenic non-human animals capable of producing human antibodies as described in US patent 5,545,806. Also fragments derived from these monoclonal antibodies such as Fab, $F(ab)'_2$ and scFv ("single chain variable fragment"), providing they have retained the original binding properties, form part of the present invention. Such fragments are commonly generated by, for instance, enzymatic digestion of the antibodies with papain, pepsin, or other proteases. It is well known to the person skilled in the art that monoclonal antibodies, or fragments thereof, can be modified for various uses. The antibodies involved in the invention can be labeled by an appropriate label of the enzymatic, fluorescent, or radioactive type.

Also within the scope of the invention are oligoribonucleotide sequences, that include anti-sense RNA and DNA molecules and ribozymes that function to inhibit the translation of Hep2 mRNA. Anti-sense RNA and DNA molecules act to directly block the translation of mRNA by binding to targeted mRNA and preventing protein translation. In regard to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between -10 and +10 regions of the Hep2 nucleotide sequence, are preferred.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridization of the ribozyme molecule to complementary target RNA, followed by a endonucleolytic cleavage. Within the scope of the invention are engineered hammerhead motif ribozyme molecules that specifically and efficiently catalyze endonucleolytic cleavage of Hep2 RNA sequences. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the

region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays.

Both anti-sense RNA and DNA molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors that incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize anti-sense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

In another embodiment of the invention the above-described molecules can be used as a medicament for treatment of diseases as described herein. It should be clear that the therapeutic method of the present invention against for example, bot not limited to this, the prevention of tumour invasion can also be used in combination with any other tumour therapy known in the art such as irradiation, chemotherapy or surgery.

The term 'medicament to treat' relates to a composition comprising molecules as described above and a pharmaceutically acceptable carrier or excipient (both terms can be used interchangeably) to treat diseases as indicated above.

The administration of a compound, an antagonist or agonist of the Hep2 or a pharmaceutically acceptable salt thereof may be by way of oral, inhaled or parenteral administration. The active compound may be administered alone or preferably formulated as a pharmaceutical composition.

An amount effective to treat the disorders hereinbefore described depends on the usual factors such as the nature and severity of the disorders being treated and the weight of the mammal. However, a unit dose will normally contain 0.01 to 50 mg for example 0.01 to 10 mg, or 0.05 to 2 mg of Hep2 agonist or antagonist or a pharmaceutically acceptable salt thereof. Unit doses will normally be administered once or more than once a day, for example 2, 3, or 4 times a day, more usually 1 to 3 times a day, such that the total daily dose is normally in the range of 0.0001 to 1

mg/kg; thus a suitable total daily dose for a 70 kg adult is 0.01 to 50 mg, for example 0.01 to 10 mg or more usually 0.05 to 10 mg.

It is greatly preferred that the compound or a pharmaceutically acceptable salt thereof is administered in the form of a unit-dose composition, such as a unit dose oral, parenteral, or inhaled composition.

Such compositions are prepared by admixture and are suitably adapted for oral, inhaled or parenteral administration, and as such may be in the form of tablets, capsules, oral liquid preparations, powders, granules, lozenges, reconstitutable powders, injectable and infusable solutions or suspensions or suppositories or aerosols.

Tablets and capsules for oral administration are usually presented in a unit dose, and contain conventional excipients such as binding agents, fillers, diluents, tabletting agents, lubricants, disintegrants, colourants, flavourings, and wetting agents. The tablets may be coated according to well-known methods in the art.

Suitable fillers for use include cellulose, mannitol, lactose and other similar agents. Suitable disintegrants include starch, polyvinylpyrrolidone and starch derivatives such as sodium starch glycollate. Suitable lubricants include, for example, magnesium stearate. Suitable pharmaceutically acceptable wetting agents include sodium lauryl sulphate.

These solid oral compositions may be prepared by conventional methods of blending, filling, tabletting or the like. Repeated blending operations may be used to distribute the active agent throughout those compositions employing large quantities of fillers. Such operations are, of course, conventional in the art.

Oral liquid preparations may be in the form of, for example, aqueous or oily suspensions, solutions, emulsions, syrups, or elixirs, or may be presented as a dry product for reconstitution with water or other suitable vehicle before use. Such liquid preparations may contain conventional additives such as suspending agents, for example sorbitol, syrup, methyl cellulose, gelatin, hydroxyethylcellulose, carboxymethyl cellulose, aluminium stearate gel or hydrogenated edible fats, emulsifying agents, for example lecithin, sorbitan monooleate, or acacia; non-aqueous vehicles (which may include edible oils), for example, almond oil, fractionated coconut oil, oily esters such as esters of glycerine, propylene glycol, or ethyl alcohol; preservatives, for example methyl or propyl p-hydroxybenzoate or sorbic acid, and if desired conventional flavouring or colouring agents.

Oral formulations also include conventional sustained release formulations, such as tablets or granules having an enteric coating.

Preferably, compositions for inhalation are presented for administration to the respiratory tract as a snuff or an aerosol or solution for a nebulizer, or as a microfine powder for insufflation, alone or in combination with an inert carrier such as lactose. In such a case the particles of active compound suitably have diameters of less than 50 microns, preferably less than 10 microns, for example between 1 and 5 microns, such as between 2 and 5 microns. A favoured inhaled dose will be in the range of 0.05 to 2 mg, for example 0.05 to 0.5 mg, 0.1 to 1 mg or 0.5 to 2 mg.

For parenteral administration, fluid unit dose forms are prepared containing a compound of the present invention and a sterile vehicle. The active compound, depending on the vehicle and the concentration, can be either suspended or dissolved. Parenteral solutions are normally prepared by dissolving the compound in a vehicle and filter sterilising before filling into a suitable vial or ampoule and sealing. Advantageously, adjuvants such as a local anaesthetic, preservatives and buffering agents are also dissolved in the vehicle. To enhance the stability, the composition can be frozen after filling into the vial and the water removed under vacuum. Parenteral suspensions are prepared in substantially the same manner except that the compound is suspended in the vehicle instead of being dissolved and sterilised by exposure to ethylene oxide before suspending in the sterile vehicle. Advantageously, a surfactant or wetting agent is included in the composition to facilitate uniform distribution of the active compound. Where appropriate, small amounts of bronchodilators for example sympathomimetic amines such as isoprenaline, isoetharine, salbutamol, phenylephrine and ephedrine; xanthine derivatives such as theophylline and aminophylline and corticosteroids such as prednisolone and adrenal stimulants such as ACTH may be included.

As is common practice, the compositions will usually be accompanied by written or printed directions for use in the medical treatment concerned.

The present invention further provides a pharmaceutical composition for use in the treatment and/or prophylaxis of herein described disorders which comprises a molecule or a pharmaceutically acceptable salt thereof, or a pharmaceutically acceptable solvate thereof, and, if required, a pharmaceutically acceptable carrier thereof. Another pharmaceutically acceptable composition is an inhalation

composition, suitably in unit dosage form. Such compositions may be prepared in the manner as hereinbefore described.

Genetic constructs of Hep2 or functional fragments thereof can also be used in gene therapy. A 'genetic construct' means that the coding information of Hep2, as depicted in SEQ ID No:2, or a functional fragment thereof is operably linked to elements known in the art that can provide transcription of the Hep2, such as a promoter and/or enhancer sequence. Gene therapy means the treatment by the delivery of therapeutic nucleic acids to patient's cells. This is extensively reviewed in Lever and Goodfellow 1995; *Br. Med Bull.*, 51, 1-242; Culver 1995; Ledley, F.D. 1995. *Hum. Gene Ther.* 6, 1129. To achieve gene therapy there must be a method of delivering genes to the patient's cells and additional methods to ensure the effective production of any therapeutic genes. There are two general approaches to achieve gene delivery; these are non-viral delivery and virus-mediated gene delivery. As an example, but not limited to this, is the use of a virus-mediated gene delivery system with replication defective retroviruses to stably introduce genes into patient's cells.

In another embodiment of the invention polymorphisms in the Hep2 gene can be detected and used diagnostically to identify patients at risk to develop diseases as herein described. Polymorphisms can occur in regulatory regions of the Hep2 gene, for example in 5' and 3' untranslated regions. Said polymorphism can induce a higher or lower expression level of Hep2. Alternatively polymorphisms can be found in the coding region of Hep2 and can lead to less susceptibility to physiological regulation or to a lower or higher activity of Hep2.

In yet another specific embodiment the invention may relate to a method of diagnosis. As an example, but not limited to this, tumour tissue can be diagnosed and it can be predicted if said tumour cells have invasive and metastatic properties. In said type of diagnostic method the essential steps comprise of taking a biopsy of the tumour cells to be diagnosed, determining the original cell type, extracting total proteins of the tumour cells and measuring the amount and/or catalytic activity of heparanase2 in the presence of a suitable heparanase substrate. Said amount of Hep2 catalytic activity is compared with a reference sample, originating from normal cells of the same type, preferentially from the same individual.

In yet another specific embodiment, the Hep2 protein or a functional fragment thereof can be used for heparan sulphate sequencing.

Finally, in a last embodiment, the Hep 2 protein or a functional fragment thereof can be used the preparation of low molecular weight heparin or heparan sulphate.

The present invention will now be illustrated by reference to the following examples which set forth particularly advantageous embodiments. However, it should be noted that these embodiments are illustrative and cannot be construed as to restrict the invention in any way.

Examples

Cloning of hep2

A heparanase homologue was identified by BLAST - searching the human EST-database, querying with the human heparanase sequence (Vlodavsky *et al.*, (1999), *Nat. Med.* 5, 793). One EST-entry (GenBank ID: AI222323; corresponding to the partial sequence of IMAGE clone 1843155 and containing an open reading frame (ORF) of 117 nucleotides followed by a STOP codon) was identified as encoding a C-terminal peptide of 39 amino acids (aa) with moderate but possibly significant sequence similarity to the C-terminal end of heparanase (41% identities, 56% positives). Sequencing of IMAGE clone 1843155 verified this homology, and resulted in additional 5' sequence information, yielding an ORF encoding a peptide of 71 aa, showing 40% sequence identity and 50% sequence similarity to the C-terminal end of heparanase (Vlodavsky *et al.*, (1999), *Nat. Med.* 5, 793). Yet, the presence of upstream STOP codons in the corresponding reading frame suggested that IMAGE clone 1843155 represents the copy of an unspliced or illegitimate transcript, or a cloning artefact (in other words was a 'chimeric' cDNA).

Based on this sequence information, primer pairs were designed to amplify the part of the IMAGE clone 1843155 cDNA that encoded a peptide with significant homology to heparanase (Vlodavsky *et al.*, (1999), *Nat. Med.* 5, 793), corresponding to residues 1680 - 1830 of SEQ ID NO :1. This PCR-probe was used for the analysis of human Multiple Tissue Northern Blots (Clontech), and identified transcripts of approximately 5.4 kb in several tissues, notably in heart, skeletal muscle, pancreas, and to a lesser extent in liver, kidney and lung. Whereas heparanase appears to be encoded by transcripts of 4.4 and 2.0 kb (Hulett et al., 1999), this result confirmed that a distinct transcript encoding peptide related (but not identical) to heparanase was present in several polyA$^+$-RNA preparations from different origins.

Therefore, additional primers were designed for RACE experiments and the isolation of the whole coding cDNA for what is designated as heparanase 2.

Heparanase 2 activity

Heparanase activity-assays were performed on different transfected cell lines (293, Hela, Caco 2, MDCK, COS 1, CHO-K1), at various pH values, ranging from pH 2.5 to 7.5. The cells were transfected with the respective heparanase-1 (Vlodavsky I. *et al.*, (1999) *Nat. Med.* 5, 793) or -2 expression-vectors, or empty vectors as controls. After 48h, the cells were lysed and the extracts were incubated with radiosulfate-labeled HS-substrates, for 18-24h. Reactions were stopped by boiling the samples in reducing sample buffer and analyzed in SDS-PAGE, followed by autoradiography. In a first assay (in which protein-free, sulfate-labelled, single HS-chains were used as substrate) all extracts from cell lines that were transfected with heparanase-1 (figuring as positive control) completely degraded the HS chains into small labelled fragments. Heparanase-1 activity was optimal at pH 3.5-6.0. For cells transfected with the various splice variants of heparanase-2, in contrast, HS-degrading activity could only be demonstrated in Hela cells. In these cells, and with this particular substrate and detection assay, heparanase-2 activity was only apparent as a discrete broadening of the labelled HS band towards the lower molecular weight region. Unlike for heparanase-1, the optimal pH value for heparanase-2 activity was observed to be situated in the range of pH 5.0-6.0. The sulfate-labelled HS-fragments generated by heparanase-2 (and its splice variants herein described before) were clearly much larger than those created by heparanase-1, indicating that heparanase-2 (and its splice variants herein described before) recognizes and cleaves sites within the HS chains that are different from and much less abundant than those that are recognized and cleaved by heparanase-1. It is well known that sulfation does not occur homogeneously along the length of the HS chains, and that the parts of the HS-chains that are close to the linkage region of the chain (the segment that joins the chains to the core proteins of the proteoglycans) are areas of low modification. In a second assay, designed to test whether heparanase-2 might preferentially cleave these protein-proximal and low-sulfated areas of the HS chains (while respecting the more sulfated regions that are degraded by heparanase-1), we incubated extracts of heparanase-1 or heparanase-2 transfected cells with HS-chain clusters. In this form of substrate, three HS-chains are still connected to each other via linkage (by their

reducing ends) to a short protease-resistant peptide (originating from the core protein). A clear reduction is observed of the sizes of these HS-clusters in heparanase-2-transfected cells, which is to be expected if heparanase-2 cleaves HS within or close to the HS-protein linkage region. To further confirm these results, we also tagged protein-free single HS chains (prepared from proteoglycan by alkaline treatment) with a fluorochrome at their reducing ends (aldehyde coupling to APTS or AMAC). Incubation of these end-labelled HS-chains with extracts from heparanase-2-transfected cells resulted in small fluorochrome-labeled fragments. All together, these results clearly demonstrate that heparanase-2 (and its splice variants thereof) cleaves HS, at specific sites that most often occur close to the linkage region. In this sense, the activity of heparanase-2 (and its splice variants thereof) is clearly distinct of that of heparanase-1.

Expression of heparanase-2

The comparison between two different heparanase-2AB expression-constructs is currently being tested. The first construct (construct-1) contains both the first and the second ATG codon-sequences present in SEQ ID NO: 1. Translation from the first ATG can produce the hep-2 polypeptide as described in WO 01/21814, (which has an N-terminal extension of 10 amino acids (MRVLCAFPEA) as compared to SEQ ID NO: 2) or it can produce the hep-2 polypeptide claimed in the present invention (SEQ ID NO: 2). In a second construct (construct-2), only the second ATG is retained (by changing the first ATG into GTG). After transfection with construct-1, several clones were selected, based on the expression of similar levels of hep-2 mRNA. Clones transfected with construct-1 (with both ATGs present) express heparanase-2 protein and show enhanced heparanase activity (as demonstrated in our experiments, see below). However, it is expected that cells transfected with construct-2 (removal of the 1st ATG), will express much higher levels of a hep-2 protein which can be detected by heparanase-2-specific antibody in immunoblots, and of much higher heparanase activity. We postulate that the nucleotide sequences upstream of the second ATG are not needed for encoding protein and activity, and also reduce the translational efficiency. We also postulate that a heparanase-2AB protein with an N-terminal extention of 10 aminoacids (MRVLCAFPEA) has only a very low activity or no activity and that the correct amino acid sequence of catalytically active hep-2 is depicted in SEQ ID NO: 2. We believe this is the main reason why catalytically active hep-2 can

be detected in the present invention and not in WO 01/21814. Thus, extraneous amino-acid sequences at the N-terminus of the hep-2 protein as depicted in SEQ ID NO: 2 can result in non-functional protein.

Generation of anti-heparanase-2 antibodies

Based on the primary protein sequence of heparanase-2AB, two potentially antigenic peptide sequences were selected which are present in all heparanase-2 splice variants, but not in heparanase-1. Corresponding synthetic peptides (C)+QNLRNPAKSRGGPGP (aa 116-130 of SEQ ID NO: 2) and GLQRKPRPGRVIRDK +(C) (aa 453-467 of SEQ ID NO: 2) were synthesized, with an additional cysteine (at the NH_2- or the COOH-terminus, as indicated) to allow specific coupling to KLH (Keyhole Limpet Hemacyanin), used as carrier protein. Two rabbits were immunized with the peptide mixtures, according to standard immunization protocols. Blood samples were taken, 10 days after every boost immunization. After the final bleeding, specific antibodies were isolated by affinity-purification on corresponding peptides coupled to EAH-Sepharose (wash conditions: PBS; elution conditions: 100mM Glycine pH 2.5), and stored in PBS pH 7.2, 0.01% sodium azide, 1% BSA. The anti-peptide antibodies react specifically with all the four splice variants of heparanase-2, in immunoblots of cells that were transfected with the different heparanase-2 splice variants, clearly revealing the size differences of the different splice variants.

Tables

Table 1:

	Sources		Mw
Hep2AB	Heart library	RACE	65411
Hep2A	CaCo2, Heart library, skeletal muscle library	PCR, RACE	58894
Hep2B	Skeletal muscle library	PCR, RACE	59268
Hep2–	Heart library	PCR, RACE	52752

Table 2:

No	Marathon ready libraries (Clontech)	5´RACE			Nested 5´RACE		
4	Heart	AP1	PR 22-04	SEQ ID No: 9	AP2	PR 22-13	SEQ ID No: 12
4a	Heart	AP1	PR 23-24	SEQ ID No: 15	AP2	PR 23-26	SEQ ID No: 16
15	Skeletal muscle	AP1	PR 23-24	SEQ ID No: 15	AP2	PR 23-26	SEQ ID No: 16
22	Skeletal muscle	AP1	PR 22-91	SEQ ID No: 14	AP2	PR 23-24	SEQ ID No: 15
32	Skeletal muscle	AP1	PR 23-48	SEQ ID No: 18	AP2	PR 23-50	SEQ ID No: 19
40	Heart	AP1	PR 23-48	SEQ ID No: 18	AP2	PR 23-50	SEQ ID No: 19

Table 3:

LabID:	Sequence:	Orientation	Length	SEQ ID
PR22-04	AGGCCAAAGCATTGACATTCTTGACCA	AS	30	9
PR22-10	AATGGCCAGCCCTTAGTGATGGTGG	S	25	10
PR22-11	TCGGTAGCGGCAGGCCAAAGCATTG	AS	25	11
PR22-13	TCCCGTCGTCCACCATCACTAAGGG	AS	25	12
PR22-90	CTTACAGATTCATGAAGGTGGCAGG	S	25	13
PR22-91	CTGATTGGCCAGCATTCCTAAAGTG	AS	25	14
PR23-24	TGTAGCAATGTTGCCAGGTAAGTGC	AS	25	15
PR23-26	ACTTCCTGCCACCTTCATGAATCTG	AS	25	16
PR23-41	TGGCTCGATTTCCTAAGCTCCAAGC	S	25	17
PR23-48	GCAACAGGCTCTTCAGCTGGATGTA	AS	25	18
PR23-50	ATGTAATCCTTTCCCAACTGGCTGC	AS	25	19
AP1	CCATCCTAATACGACTCACTATAGGGC		27	20
AP2	ACTCACTATAGGGCTCGAGCGGC		23	21

Table 4:

	Multiple tissue Northerns (Clontech)																			
	Adult Human H2								Adult Human H								Fetal Human			
	Spleen	Thymus	Prostate	Testis	Ovary	Small intestine	Colon	Periph. Blood Leuc	Heart	Brain	Placenta	Lung	Liver	Skel. Muscle	Kidney	Pancreas	Brain	Lung	Liver	kidney
Hep1*	+ + +	-						+++	-	-	+ + +	-	-	-	-	-			+ + +	
Hep2	-	-	-	-	-	-	-	-	+ + +	-	-	-	+	+ + +	+	+ + +	-	+	-	-

*data taken from Hulett *et al.*, (1999), *Nat. Med.* 5, 803

Materials and Methods

Ligations, transformations of bacteria, and other standard molecular biological techniques were performed according to established procedures (Sambrook et al., 1989).

Nucleotide sequence analysis:

Nucleotide sequence analysis was performed using the dideoxy-mediated chain termination procedure with fluorescent primers and the Thermo Sequenase Primer Cycle Sequencing Kit (Amersham-Pharmacia). Approximately 1 µg of plasmid DNA was used as template. The reactions were performed using a Gene-amp 9600 (Perkin-Elmer) with following cycling parameters: incubation for 2.5 min at 95°C, 20 cycles of respectively 30 sec at 95°C, 30 sec at 55°C, and 1 min at 72°C. This was followed by 10 cycles of respectively 30 sec at 95°C, 1 min at 72°C. The cycle sequencing reactions were analysed by electrophoresis using an A.L.F. DNA Sequencer (Amersham-Pharmacia) on standard 30 cm, 6 % Hydrolink Long Ranger gels (AT Biochem). Sequence-analysis was performed using bio-informatic tools.

Polymerase chain reaction:

All oligonucleotides were purchased from Eurogentec. PCR conditions used were identical to the ones used for RACE (see there).

Labelling of DNA probes:

Molecular probes used in this study were labelled with ^{32}P-dCTP (NEN-Dupont) using Ready-to-Go oligolabelling beads (Pharmacia), according to the manufacturer's instructions.

RACE (Rapid Amplification of cDNA Ends):

RACE was performed on a library of adaptor-ligated foetal brain cDNA, using the conditions stipulated and reagents (Marathon cDNA Amplification kit) provided by the supplier (Clontech, Palo Alto, CA). The cDNAs were amplified through a two-step PCR protocol. The first PCR used a gene-specific and an anchor primer provided by the supplier. Then 5 µl of the first PCR reaction was used as template for the second PCR-reaction, using a second gene-specific nested primer and a nested anchor primer provided by the supplier. The products of the second PCR were analysed by electrophoresis in a 1 % agarose gel. Distinct PCR-products were gel purified using either the Wizzard DNA clean-up system (Promega, Madison, WI) or Qiaquick (Qiagen Inc., Santa Clarita, CA). PCR-products were T/A cloned using the vector pGEMTeasy. Multiple independent RACE clones were sequenced in each case.

Primers AP1 (5′-CCATCCTAATACGACTCACTATAGGGC-3′) and AP2 (5′-ACTCACTATAGGGCTCGAGCGGC-3′) were contained in the Marathon ready cDNA kit of Clontech

Thermal cycling for primary and secondary (nested) PCR and touchdown PCR was performed in a Perkin Elmer Thermal Cycler 480, with the following parameters

- 94 C for 1 min
- 5 cycles:

94 C 30 sec

72 C 4 min

- 5 cycles:

94 C 30 sec

70 C 4 min

- 25 cycles:

94 C 30 sec

68 C 4 min

Northern blotting:

Premade Northern blots of poly(A) RNA from multiple human tissues were obtained from Clontech. Hybridization was performed with ^{32}P-oligolabeled probes for two hours at 68°C, using Expresshyb solution (Clontech) according to the manufacturer's specifications. Dehybridisation included washing at room temperature for 30 min with 2.0 % SSC, 0.05 % SDS and a high stringency wash for 30 min with 0.1 % SSC, 0.1 % SDS at 65°C.

Isolation of HS-chains

10^9 human lung fibroblasts were grown to confluency and incubated for 48h with 10µCi of 35S H_2SO_4/ml growth medium (Ham's F12, 10%FCS, 0,1mM Mg_2SO_4). The medium was collected and cells were washed 3x with PBS- and scraped in extraction buffer (150mM NaCl, 20 mM Tris-HCl pH 8.0, 0.5% Triton-X 100(Ultrapure, Boehringer), 50mM 6-aminohexanoic acid, 10mM EDTA, 5mM benzamidine, 5mM NEM, 1µg/ml Pepstatin A). The extract and the medium were applied to a DEAE column (5ml packed beads). Columns were washed with 5 column volumes TUT buffer (6M Urea, 50mM Tris pH 8.0, 0.5% Triton X-100) containing 200mM NaCl, followed by 2 column volumes of TUT-buffer, and eluted by 2 column volumes 1M NaCl in TUT buffer. The eluant was adjusted to 200mM NaCl with TUT-buffer and reconcentrated on a DEAE-column (250µl packed beads) equilibrated with 200mM NaCl, 20mM Tris pH 7.4, 0.1% Triton X 100. The column was eluted with 400µl 1M NaCl, 20mM Tris pH 7.4, 0.1% Triton X 100 and fractions of 100µl were collected and precipitated by addition of 10µg glycogen and 3 volumes EtOH and incubated at –20°C. Precipitates were pelleted by 10 min centrifugation at 14 000 rpm, 4°C. After washing the pellets with 70% EtOH, samples were resuspended in 105µl 100mM NaCl, 20mM Tris pH 8.0, Triton-X 100. An equal volume of 1M KOH was added for alkaline treatment and samples were incubated at 4°C for 12h. The reaction was stopped by neutralization with 7µl of acetic acid. 2ml of 20mM Tris pH 8.0, Triton-X 100 were added and the samples concentrated again on DEAE (150 µl packed beads) columns and precipitated with EtOH, as above. The precipitates were pelleted and resuspended in 210µl H_2O.

To remove contaminating CS-chains, an equal volume of 2xABCase buffer (100mM NaCl, 100mM NaAcetate,100mM Tris pH 8.3) was added to the samples prior to treatment with 25mU chondroitinase ABC for 3h with the addition of 25mU fresh enzyme after 1.5h. Reactions were stopped by EtOH precipitation (see above). Samples were either resuspended in phosphate buffer and used directly for heparanase activity assays, or were subject to end labeling with fluorochromes (see below).

Generation of HS-chain clusters

HS-chain clusters were prepared from human lung fibroblasts proteoglycans by treating the 1st DEAE eluant (see above) with 500µl proteinase K (2mg/ml in 20mM Tris pH 8.0) for 1h at 55 °C. The reaction was stopped by addition of 100µl 100mM PMSF. The digest was adjusted to 200mM NaCl with TUT-buffer and concentrated on a DEAE-column (200µl packed beads) equilibrated and washed with 200mM NaCl, 20mM Tris pH 7.4, 0.1% Triton X 100. The column was eluted with 400µl 1M NaCl, 20mM Tris pH 7.4, 0.1% Triton X 100 and fractions of 100µl were collected and precipitated by addition of 10µg glycogen and 3 volumes EtOH and incubated at –20°C. Precipitates were pelleted by 10 min centrifugation at 14 000 rpm, 4°C. After washing the pellets with 70% EtOH, samples were resuspended in 1xABCase buffer and treated with ABCase as described above to remove CS-contaminants.

End labeling of HS-chains with fluorochromes

Single HS-chains derived from alkaline treatment of HS-proteoglycans (see above) were end-labeled by aldehyde-coupling of either 2-AMAC (2-aminoacridone) or APTS (8-aminopyrene-1,3,6-trisulfonic acid) to the chains. The HS-chains were either incubated with 100mM APTS or 50mM 2-AMAC, both in formamide, by incubation with 1M NaCNBH3 at 37°C for 12h. Reaction volume 2µl. The reaction was stopped by addition of 100 µl 20mM Tris pH 7.4. 2-AMAC-labeled chains were applied to DEAE-columns. Free 2-AMAC was eluted from the column by washing with 200mM NaCl in TUT buffer and labeled chains were eluted with 1M NaCl. The eluant was precipitated with EtOH. APTS-labeled chains were directly purified by EtOH precipitation. The precipitates were dissolved in phosphate buffer for use in heparanase activity assays.

Heparanase assay

Cells were grown to 90% confluency in 6-well plates and transfected with 1µg of the respective heparanase expression- or control vectors, using Fugene (Boehringer) according to the manufacturer's instructions. After 48h cells were washed and lysed in 250µl lysis buffer (PBS^{++} pH 7.2, 1% NP40, protease inhibitors) per well for 30min at 4°C. The lysate was cleared by centrifugation and 50µl of the supernatant was incubated at the given pH-values with the respective radiolabeled HS-samples at 37°C, ON. Reactions were stopped by adding sample buffer and 5 min of boiling of the samples before they were analyzed in SDS-PAGE. HS-chains and degradation products were visualized by autoradiography or by UV-detection in the case of fluorochrome labeled chains.

Cells and cell culture

COS-1 (SV40 transformed, african green monkey kidney, CRL-1650), 293 (human embryonal kidney, CRL-1573), CaCo2 (human colon adenocarcinoma, HTB-37), CHO-K1 (Chinese hamster ovary, CCL-61), HeLa (cervix carcinoma, CCL-2), MDCK II (dog kidney, CCL-34) were grown in DMEM/F12 (Gibco/Brl), 10% FCS and kept under a humid 5% CO_2 atmosphere. Cells were split 1:5 on a routine basis once they reached 90% confluency.

Claims

1. An isolated polypeptide having the primary structural information of amino acids 1-582 as set forth in SEQ ID NO: 2 or any functional fragment or allelic variant or splice variant thereof possessing the biological property of having heparanase catalytic activity.

2. An isolated polynucleotide encoding the polypeptide of claim 1 and set forth in SEQ ID NO: 1 or any functional fragment or allelic variant or splice variant thereof possessing the biological property of encoding heparanase catalytic activity.

3. A polypeptide according to claim 1 wherein said polypeptide shares at least 70% homology with SEQ ID NO: 2.

4. A polynucleotide according to claim 2 wherein said polynucleotide shares at least 70 % homology with SEQ ID NO: 1.

5. A vector comprising the nucleic acid molecule of claims 2 or 4.

6. The vector of claim 5, said vector being an expression vector, in particular further comprising a regulatory element.

7. A genetically engineered host cell comprising the expression vector of claim 6.

8. A recombinant protein comprising a functional fragment or a splice variant of SEQ ID NO: 2

9. A method to screen for molecules with potential to antagonize or agonize the heparanase 2 catalytic activity according to claim 1 comprising:

a. Adding a molecule to heparanase 2 enzyme in the presence of a heparanase substrate and

b. Evaluating the effect of the molecule on the catalytic activity of said heparanase 2 enzyme towards said heparanase substrate

10. A molecule obtained in the screening assay of claim 9.

11. A molecule according to claim 10 for use as a medicament.

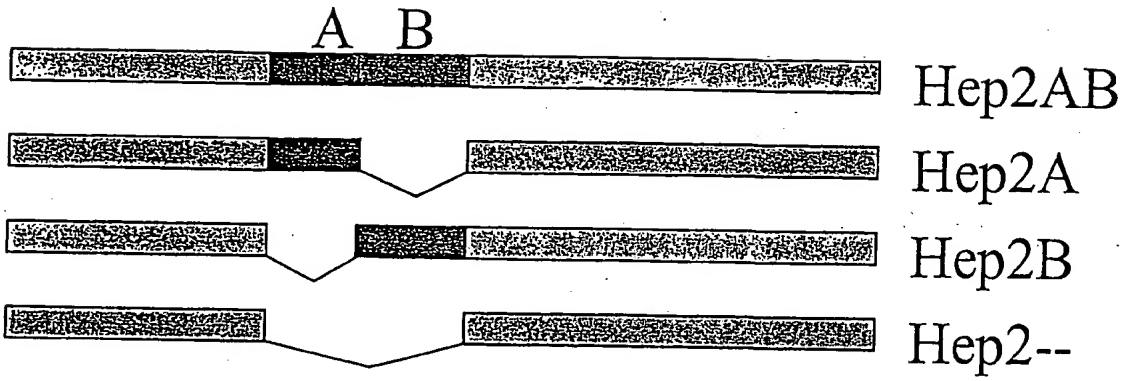
12. Use of a polynucleotide according to claim 2 for use in gene therapy.

13. Use of a polypeptide according to claim 1 as a medicament.

14. An antibody which specifically recognizes and binds to a polypeptide as defined in claim 1.

15. Use of polymorphisms in the polynucleotide sequence according to claim 2 to identify individuals having a predisposition to acquire diseases resulting from a shortage or excessive activity of a polypeptide according to claim 1.

Figure 1



SEQUENCE LISTING

<110> Vlaams Interuniversitair Instituut voor Biotechnol

<120> A second human heparanase, and splice variants thereof, with a predominant expression in skeletal muscle, heart and pancreas.

<130> GDA/HEP/V070

<140>
<141>

<150> 00202442.0
<151> 2000-07-12

<160> 20

<170> PatentIn Ver. 2.1

<210> 1
<211> 1847
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (85)..(1833)

<400> 1
aattagtcac tagcgagacc cagtaggaag agaggtttaa atcagaggga ttgaatgagg 60

gtgctttgtg ccttccctga agcc atg ccc tcc agc aac tcc cgc ccc ccc 111
Met Pro Ser Ser Asn Ser Arg Pro Pro
1 5

gcg tgc cta gcc ccg ggg gct ctc tac ttg gct ctg ttg ctc cat ctc 159
Ala Cys Leu Ala Pro Gly Ala Leu Tyr Leu Ala Leu Leu Leu His Leu
10 15 20 25

tcc ctt tcc tcc cag gct gga gac agg aga ccc ttg cct gta gac aga 207
Ser Leu Ser Ser Gln Ala Gly Asp Arg Arg Pro Leu Pro Val Asp Arg
30 35 40

gct gca ggt ttg aag gaa aag acc ctg att cta ctt gat gtg agc acc 255
Ala Ala Gly Leu Lys Glu Lys Thr Leu Ile Leu Leu Asp Val Ser Thr
45 50 55

```
aag aac cca gtc agg aca gtc aat gag aac ttc ctc tct ctg cag ctg   303
Lys Asn Pro Val Arg Thr Val Asn Glu Asn Phe Leu Ser Leu Gln Leu
        60                  65                  70

gat ccg tcc atc att cat gat ggc tgg ctc gat ttc cta agc tcc aag   351
Asp Pro Ser Ile Ile His Asp Gly Trp Leu Asp Phe Leu Ser Ser Lys
    75                  80                  85

cgc ttg gtg acc ctg gcc cgg gga ctt tcg ccc gcc ttt ctg cgc ttc   399
Arg Leu Val Thr Leu Ala Arg Gly Leu Ser Pro Ala Phe Leu Arg Phe
90                  95                  100                 105

ggg ggc aaa agg gcc gac ttc ctg cag ttc cag aac ctg agg aac ccg   447
Gly Gly Lys Arg Ala Asp Phe Leu Gln Phe Gln Asn Leu Arg Asn Pro
                110                 115                 120

gcg aaa agc cgc ggg ggc ccg ggc ccg gat tac tat ctc aaa aac tat   495
Ala Lys Ser Arg Gly Gly Pro Gly Pro Asp Tyr Tyr Leu Lys Asn Tyr
            125                 130                 135

gag gat gac att gtc cga agt gat gtt gcc tta gat aaa cag aaa ggc   543
Glu Asp Asp Ile Val Arg Ser Asp Val Ala Leu Asp Lys Gln Lys Gly
        140                 145                 150

tgc aag att gcc cgg cac cct gat gtt atg ctg gag ctc caa agg gag   591
Cys Lys Ile Ala Arg His Pro Asp Val Met Leu Glu Leu Gln Arg Glu
    155                 160                 165

aag gca gct cag atg cat ctg gtt ctt cta aag gag caa ttc tcc aat   639
Lys Ala Ala Gln Met His Leu Val Leu Leu Lys Glu Gln Phe Ser Asn
170                 175                 180                 185

act aac agt aat ctc ata tta aca gcc agg tct cta gac aaa ctt tat   687
Thr Asn Ser Asn Leu Ile Leu Thr Ala Arg Ser Leu Asp Lys Leu Tyr
                190                 195                 200

aac ttt gct gat tgc tct gga ctc cac ctg ata ttt gct cta aat gca   735
Asn Phe Ala Asp Cys Ser Gly Leu His Leu Ile Phe Ala Leu Asn Ala
            205                 210                 215

ctg cgt cgt aat ccc aat aac tcc tgg aac agt tct agt gcc ctg agt   783
Leu Arg Arg Asn Pro Asn Asn Ser Trp Asn Ser Ser Ser Ala Leu Ser
        220                 225                 230

ctg ttg aag tac agc gcc agc aaa aag tac aac att tct tgg gaa ctg   831
Leu Leu Lys Tyr Ser Ala Ser Lys Lys Tyr Asn Ile Ser Trp Glu Leu
    235                 240                 245
```

```
ggt aat gag cca aat aac tat cgg acc atg cat ggc cgg gca gta aat   879
Gly Asn Glu Pro Asn Asn Tyr Arg Thr Met His Gly Arg Ala Val Asn
250                 255                 260                 265

ggc agc cag ttg gga aag gat tac atc cag ctg aag agc ctg ttg cag   927
Gly Ser Gln Leu Gly Lys Asp Tyr Ile Gln Leu Lys Ser Leu Leu Gln
                270                 275                 280

ccc atc cgg att tat tcc aga gcc agc tta tat ggc cct aat att ggg   975
Pro Ile Arg Ile Tyr Ser Arg Ala Ser Leu Tyr Gly Pro Asn Ile Gly
            285                 290                 295

cgg ccg agg aag aat gtc atc gcc ctc cta gat gga ttc atg aag gtg   1023
Arg Pro Arg Lys Asn Val Ile Ala Leu Leu Asp Gly Phe Met Lys Val
        300                 305                 310

gca gga agt aca gta gat gca gtt acc tgg caa cay tgc tac att gat   1071
Ala Gly Ser Thr Val Asp Ala Val Thr Trp Gln His Cys Tyr Ile Asp
    315                 320                 325

ggc cgg gtg gtc aag gtg atg gac ttc ctg aaa act cgc ctg tta gac   1119
Gly Arg Val Val Lys Val Met Asp Phe Leu Lys Thr Arg Leu Leu Asp
330                 335                 340                 345

aca ctc tct gac cag att agg aaa att cag aaa gtg gtt aat aca tac   1167
Thr Leu Ser Asp Gln Ile Arg Lys Ile Gln Lys Val Val Asn Thr Tyr
                350                 355                 360

act cca gga aag aag att tgg ctt gaa ggt gtg gtg acc acc tca gct   1215
Thr Pro Gly Lys Lys Ile Trp Leu Glu Gly Val Val Thr Thr Ser Ala
            365                 370                 375

gga ggc aca aac aat cta tcc gat tcc tat gct gca gga ttc tta tgg   1263
Gly Gly Thr Asn Asn Leu Ser Asp Ser Tyr Ala Ala Gly Phe Leu Trp
        380                 385                 390

ttg aac act tta gga atg ctg gcc aat cag ggc att gat gtc gtg ata   1311
Leu Asn Thr Leu Gly Met Leu Ala Asn Gln Gly Ile Asp Val Val Ile
    395                 400                 405

cgg cac tca ttt ttt gac cat gga tac aat cac ctc gtg gac cag aat   1359
Arg His Ser Phe Phe Asp His Gly Tyr Asn His Leu Val Asp Gln Asn
410                 415                 420                 425

ttt aac cca tta cca gac tac tgg ctc tct ctc ctc tac aag cgc ctg   1407
Phe Asn Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Tyr Lys Arg Leu
                430                 435                 440
```

```
atc ggc ccc aaa gtc ttg gct gtg cat gtg gct ggg ctc cag cgg aag   1455
Ile Gly Pro Lys Val Leu Ala Val His Val Ala Gly Leu Gln Arg Lys
            445                 450                 455

cca cgg cct ggc cga gtg atc cgg gac aaa cta agg att tat gct cac   1503
Pro Arg Pro Gly Arg Val Ile Arg Asp Lys Leu Arg Ile Tyr Ala His
        460                 465                 470

tgc aca aac cac cac aac cac aac tac gtt cgt ggg tcc att aca ctt   1551
Cys Thr Asn His His Asn His Asn Tyr Val Arg Gly Ser Ile Thr Leu
    475                 480                 485

ttt atc atc aac ttg cat cga tca aga aag aaa atc aag ctg gct ggg   1599
Phe Ile Ile Asn Leu His Arg Ser Arg Lys Lys Ile Lys Leu Ala Gly
490                 495                 500                 505

act ctc aga gac aag ctg gtt cac cag tac ctg ctg cag ccc tat ggg   1647
Thr Leu Arg Asp Lys Leu Val His Gln Tyr Leu Leu Gln Pro Tyr Gly
                510                 515                 520

cag gag ggc cta aag tcc aag tca gtg caa ctg aat ggc cag ccc tta   1695
Gln Glu Gly Leu Lys Ser Lys Ser Val Gln Leu Asn Gly Gln Pro Leu
            525                 530                 535

gtg atg gtg gac gac ggg acc ctc cca gaa ttg aag ccc cgc ccc ctt   1743
Val Met Val Asp Asp Gly Thr Leu Pro Glu Leu Lys Pro Arg Pro Leu
        540                 545                 550

cgg gcc ggc cgg aca ttg gtc atc cct cca gtc acc atg ggc ttt ttt   1791
Arg Ala Gly Arg Thr Leu Val Ile Pro Pro Val Thr Met Gly Phe Phe
    555                 560                 565

gtg gtc aag aat gtc aat gct ttg gcc tgc cgc tac cga taa           1833
Val Val Lys Asn Val Asn Ala Leu Ala Cys Arg Tyr Arg
570                 575                 580

gctatcctca cact                                                  1847
```

<210> 2
<211> 582
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Pro Ser Ser Asn Ser Arg Pro Pro Ala Cys Leu Ala Pro Gly Ala
  1               5                  10                  15
Leu Tyr Leu Ala Leu Leu Leu His Leu Ser Leu Ser Ser Gln Ala Gly
```

```
            20                  25                  30
Asp Arg Arg Pro Leu Pro Val Asp Arg Ala Ala Gly Leu Lys Glu Lys
        35                  40                  45
Thr Leu Ile Leu Leu Asp Val Ser Thr Lys Asn Pro Val Arg Thr Val
    50                  55                  60
Asn Glu Asn Phe Leu Ser Leu Gln Leu Asp Pro Ser Ile Ile His Asp
65                  70                  75                  80
Gly Trp Leu Asp Phe Leu Ser Ser Lys Arg Leu Val Thr Leu Ala Arg
                85                  90                  95
Gly Leu Ser Pro Ala Phe Leu Arg Phe Gly Gly Lys Arg Ala Asp Phe
            100                 105                 110
Leu Gln Phe Gln Asn Leu Arg Asn Pro Ala Lys Ser Arg Gly Gly Pro
        115                 120                 125
Gly Pro Asp Tyr Tyr Leu Lys Asn Tyr Glu Asp Asp Ile Val Arg Ser
    130                 135                 140
Asp Val Ala Leu Asp Lys Gln Lys Gly Cys Lys Ile Ala Arg His Pro
145                 150                 155                 160
Asp Val Met Leu Glu Leu Gln Arg Glu Lys Ala Ala Gln Met His Leu
                165                 170                 175
Val Leu Leu Lys Glu Gln Phe Ser Asn Thr Asn Ser Asn Leu Ile Leu
            180                 185                 190
Thr Ala Arg Ser Leu Asp Lys Leu Tyr Asn Phe Ala Asp Cys Ser Gly
        195                 200                 205
Leu His Leu Ile Phe Ala Leu Asn Ala Leu Arg Arg Asn Pro Asn Asn
    210                 215                 220
Ser Trp Asn Ser Ser Ser Ala Leu Ser Leu Leu Lys Tyr Ser Ala Ser
225                 230                 235                 240
Lys Lys Tyr Asn Ile Ser Trp Glu Leu Gly Asn Glu Pro Asn Asn Tyr
                245                 250                 255
Arg Thr Met His Gly Arg Ala Val Asn Gly Ser Gln Leu Gly Lys Asp
            260                 265                 270
Tyr Ile Gln Leu Lys Ser Leu Leu Gln Pro Ile Arg Ile Tyr Ser Arg
        275                 280                 285
Ala Ser Leu Tyr Gly Pro Asn Ile Gly Arg Pro Arg Lys Asn Val Ile
    290                 295                 300
Ala Leu Leu Asp Gly Phe Met Lys Val Ala Gly Ser Thr Val Asp Ala
305                 310                 315                 320
Val Thr Trp Gln His Cys Tyr Ile Asp Gly Arg Val Val Lys Val Met
                325                 330                 335
Asp Phe Leu Lys Thr Arg Leu Leu Asp Thr Leu Ser Asp Gln Ile Arg
            340                 345                 350
Lys Ile Gln Lys Val Val Asn Thr Tyr Thr Pro Gly Lys Lys Ile Trp
        355                 360                 365
Leu Glu Gly Val Val Thr Thr Ser Ala Gly Gly Thr Asn Asn Leu Ser
    370                 375                 380
Asp Ser Tyr Ala Ala Gly Phe Leu Trp Leu Asn Thr Leu Gly Met Leu
385                 390                 395                 400
Ala Asn Gln Gly Ile Asp Val Val Ile Arg His Ser Phe Phe Asp His
```

405 410 415
Gly Tyr Asn His Leu Val Asp Gln Asn Phe Asn Pro Leu Pro Asp Tyr
420 425 430
Trp Leu Ser Leu Leu Tyr Lys Arg Leu Ile Gly Pro Lys Val Leu Ala
435 440 445
Val His Val Ala Gly Leu Gln Arg Lys Pro Arg Pro Gly Arg Val Ile
450 455 460
Arg Asp Lys Leu Arg Ile Tyr Ala His Cys Thr Asn His His Asn His
465 470 475 480
Asn Tyr Val Arg Gly Ser Ile Thr Leu Phe Ile Ile Asn Leu His Arg
485 490 495
Ser Arg Lys Lys Ile Lys Leu Ala Gly Thr Leu Arg Asp Lys Leu Val
500 505 510
His Gln Tyr Leu Leu Gln Pro Tyr Gly Gln Glu Gly Leu Lys Ser Lys
515 520 525
Ser Val Gln Leu Asn Gly Gln Pro Leu Val Met Val Asp Asp Gly Thr
530 535 540
Leu Pro Glu Leu Lys Pro Arg Pro Leu Arg Ala Gly Arg Thr Leu Val
545 550 555 560
Ile Pro Pro Val Thr Met Gly Phe Phe Val Val Lys Asn Val Asn Ala
565 570 575
Leu Ala Cys Arg Tyr Arg
580

<210> 3
<211> 1673
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> ()..)

<400> 3
aattagtcac tagcgagacc cagtaggaag agaggtttaa atcagaggga 50

a ttgaatgagg gtgctttgtg ccttccctga agccatgccc tccagcaact 101

cccgcccccc cgcgtgccta gccccggggg ctctctactt ggctctgttg ctccatctct 161

ccctttcctc ccaggctgga gacaggagac ccttgcctgt agacagagct gcaggtttga 221

aggaaaagac cctgattcta cttgatgtga gcaccaagaa cccagtcagg acagtcaatg 281

agaacttcct ctctctgcag ctggatccgt ccatcattca tgatggctgg ctcgatttcc 341

taagctccaa gcgcttggtg accctggccc ggggactttc gcccgccttt ctgcgcttcg 401

ggggcaaaag ggccgacttc ctgcagttcc agaacctgag gaacccggcg aaaagccgcg 461

ggggcccggg cccggattac tatctcaaaa actatgagga tgacattgtc cgaagtgatg 521

ttgccttaga taaacagaaa ggctgcaaga ttgcccggca ccctgatgtt atgctggagc 581

tccaaaggga gaaggcagct cagatgcatc tggttcttct aaaggagcaa ttctccaata 641

ctaacagtaa tctcatatta acagagccaa ataactatcg gaccatgcat ggccgggcag 701

taaatggcag ccagttggga aaggattaca tccagctgaa gagcctgttg cagcccatcc 761

ggatttattc cagagccagc ttatatggcc ctaatattgg gcggccgagg aagaatgtca 821

tcgccctcct agatggattc atgaaggtgg caggaagtac agtagatgca gttacctggc 881

aacaytgcta cattgatggc cgggtggtca aggtgatgga cttcctgaaa actcgcctgt 941

tagacacact ctctgaccag attaggaaaa ttcagaaagt ggttaataca tacactccag 1001

gaaagaagat ttggcttgaa ggtgtggtga ccacctcagc tggaggcaca aacaatctat 1061

ccgattccta tgctgcagga ttcttatggt tgaacacttt aggaatgctg gccaatcagg 1121

gcattgatgt cgtgatacgg cactcatttt ttgaccatgg atacaatcac ctcgtggacc 1181

agaattttaa cccattacca gactactggc tctctctcct ctacaagcgc ctgatcggcc 1241

ccaaagtctt ggctgtgcat gtggctgggc tccagcggaa gccacggcct ggccgagtga 1301

tccgggacaa actaaggatt tatgctcact gcacaaacca ccacaaccac aactacgttc 1361

gtgggtccat tacacttttt atcatcaact tgcatcgatc aagaaagaaa atcaagctgg 1421

ctgggactct cagagacaag ctggttcacc agtacctgct gcagccctat gggcaggagg 1481

gcctaaagtc caagtcagtg caactgaatg gccagccctt agtgatggtg gacgacggga 1541

ccctcccaga attgaagccc cgccccccttc gggccggccg gacattggtc atccctccag 1601

tcaccatggg ctttttttgtg gtcaagaatg tcaatgcttt ggcctgccgc taccgataag 1661

ctatcctcac act 1674

<210> 4
<211> 1685
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (85)..(1671)

<400> 4
aattagtcac tagcgagacc cagtaggaag agaggtttaa atcagaggga ttgaatgagg 60

gtgctttgtg ccttccctga agcc atg ccc tcc agc aac tcc cgc ccc ccc 111
Met Pro Ser Ser Asn Ser Arg Pro Pro
1 5

gcg tgc cta gcc ccg ggg gct ctc tac ttg gct ctg ttg ctc cat ctc 159
Ala Cys Leu Ala Pro Gly Ala Leu Tyr Leu Ala Leu Leu Leu His Leu
10 15 20 25

tcc ctt tcc tcc cag gct gga gac agg aga ccc ttg cct gta gac aga 207
Ser Leu Ser Ser Gln Ala Gly Asp Arg Arg Pro Leu Pro Val Asp Arg
30 35 40

gct gca ggt ttg aag gaa aag acc ctg att cta ctt gat gtg agc acc 255
Ala Ala Gly Leu Lys Glu Lys Thr Leu Ile Leu Leu Asp Val Ser Thr
45 50 55

aag aac cca gtc agg aca gtc aat gag aac ttc ctc tct ctg cag ctg 303
Lys Asn Pro Val Arg Thr Val Asn Glu Asn Phe Leu Ser Leu Gln Leu
60 65 70

gat ccg tcc atc att cat gat ggc tgg ctc gat ttc cta agc tcc aag 351
Asp Pro Ser Ile Ile His Asp Gly Trp Leu Asp Phe Leu Ser Ser Lys
75 80 85

cgc ttg gtg acc ctg gcc cgg gga ctt tcg ccc gcc ttt ctg cgc ttc 399
Arg Leu Val Thr Leu Ala Arg Gly Leu Ser Pro Ala Phe Leu Arg Phe
90 95 100 105

ggg ggc aaa agg gcc gac ttc ctg cag ttc cag aac ctg agg aac ccg 447
Gly Gly Lys Arg Ala Asp Phe Leu Gln Phe Gln Asn Leu Arg Asn Pro
110 115 120

gcg aaa agc cgc ggg ggc ccg ggc ccg gat tac tat ctc aaa aac tat 495
Ala Lys Ser Arg Gly Gly Pro Gly Pro Asp Tyr Tyr Leu Lys Asn Tyr
125 130 135

```
gag gat gcc agg tct cta gac aaa ctt tat aac ttt gct gat tgc tct   543
Glu Asp Ala Arg Ser Leu Asp Lys Leu Tyr Asn Phe Ala Asp Cys Ser
        140                 145                 150

gga ctc cac ctg ata ttt gct cta aat gca ctg cgt cgt aat ccc aat   591
Gly Leu His Leu Ile Phe Ala Leu Asn Ala Leu Arg Arg Asn Pro Asn
    155                 160                 165

aac tcc tgg aac agt tct agt gcc ctg agt ctg ttg aag tac agc gcc   639
Asn Ser Trp Asn Ser Ser Ser Ala Leu Ser Leu Leu Lys Tyr Ser Ala
170                 175                 180                 185

agc aaa aag tac aac att tct tgg gaa ctg ggt aat gag cca aat aac   687
Ser Lys Lys Tyr Asn Ile Ser Trp Glu Leu Gly Asn Glu Pro Asn Asn
                190                 195                 200

tat cgg acc atg cat ggc cgg gca gta aat ggc agc cag ttg gga aag   735
Tyr Arg Thr Met His Gly Arg Ala Val Asn Gly Ser Gln Leu Gly Lys
            205                 210                 215

gat tac atc cag ctg aag agc ctg ttg cag ccc atc cgg att tat tcc   783
Asp Tyr Ile Gln Leu Lys Ser Leu Leu Gln Pro Ile Arg Ile Tyr Ser
        220                 225                 230

aga gcc agc tta tat ggc cct aat att ggg cgg ccg agg aag aat gtc   831
Arg Ala Ser Leu Tyr Gly Pro Asn Ile Gly Arg Pro Arg Lys Asn Val
    235                 240                 245

atc gcc ctc cta gat gga ttc atg aag gtg gca gga agt aca gta gat   879
Ile Ala Leu Leu Asp Gly Phe Met Lys Val Ala Gly Ser Thr Val Asp
250                 255                 260                 265

gca gtt acc tgg caa cay tgc tac att gat ggc cgg gtg gtc aag gtg   927
Ala Val Thr Trp Gln His Cys Tyr Ile Asp Gly Arg Val Val Lys Val
                270                 275                 280

atg gac ttc ctg aaa act cgc ctg tta gac aca ctc tct gac cag att   975
Met Asp Phe Leu Lys Thr Arg Leu Leu Asp Thr Leu Ser Asp Gln Ile
            285                 290                 295

agg aaa att cag aaa gtg gtt aat aca tac act cca gga aag aag att   1023
Arg Lys Ile Gln Lys Val Val Asn Thr Tyr Thr Pro Gly Lys Lys Ile
        300                 305                 310

tgg ctt gaa ggt gtg gtg acc acc tca gct gga ggc aca aac aat cta   1071
Trp Leu Glu Gly Val Val Thr Thr Ser Ala Gly Gly Thr Asn Asn Leu
    315                 320                 325
```

```
tcc gat tcc tat gct gca gga ttc tta tgg ttg aac act tta gga atg   1119
Ser Asp Ser Tyr Ala Ala Gly Phe Leu Trp Leu Asn Thr Leu Gly Met
330                 335                 340                 345

ctg gcc aat cag ggc att gat gtc gtg ata cgg cac tca ttt ttt gac   1167
Leu Ala Asn Gln Gly Ile Asp Val Val Ile Arg His Ser Phe Phe Asp
                350                 355                 360

cat gga tac aat cac ctc gtg gac cag aat ttt aac cca tta cca gac   1215
His Gly Tyr Asn His Leu Val Asp Gln Asn Phe Asn Pro Leu Pro Asp
            365                 370                 375

tac tgg ctc tct ctc ctc tac aag cgc ctg atc ggc ccc aaa gtc ttg   1263
Tyr Trp Leu Ser Leu Leu Tyr Lys Arg Leu Ile Gly Pro Lys Val Leu
        380                 385                 390

gct gtg cat gtg gct ggg ctc cag cgg aag cca cgg cct ggc cga gtg   1311
Ala Val His Val Ala Gly Leu Gln Arg Lys Pro Arg Pro Gly Arg Val
    395                 400                 405

atc cgg gac aaa cta agg att tat gct cac tgc aca aac cac cac aac   1359
Ile Arg Asp Lys Leu Arg Ile Tyr Ala His Cys Thr Asn His His Asn
410                 415                 420                 425

cac aac tac gtt cgt ggg tcc att aca ctt ttt atc atc aac ttg cat   1407
His Asn Tyr Val Arg Gly Ser Ile Thr Leu Phe Ile Ile Asn Leu His
                430                 435                 440

cga tca aga aag aaa atc aag ctg gct ggg act ctc aga gac aag ctg   1455
Arg Ser Arg Lys Lys Ile Lys Leu Ala Gly Thr Leu Arg Asp Lys Leu
            445                 450                 455

gtt cac cag tac ctg ctg cag ccc tat ggg cag gag ggc cta aag tcc   1503
Val His Gln Tyr Leu Leu Gln Pro Tyr Gly Gln Glu Gly Leu Lys Ser
        460                 465                 470

aag tca gtg caa ctg aat ggc cag ccc tta gtg atg gtg gac gac ggg   1551
Lys Ser Val Gln Leu Asn Gly Gln Pro Leu Val Met Val Asp Asp Gly
    475                 480                 485

acc ctc cca gaa ttg aag ccc cgc ccc ctt cgg gcc ggc cgg aca ttg   1599
Thr Leu Pro Glu Leu Lys Pro Arg Pro Leu Arg Ala Gly Arg Thr Leu
490                 495                 500                 505

gtc atc cct cca gtc acc atg ggc ttt ttt gtg gtc aag aat gtc aat   1647
Val Ile Pro Pro Val Thr Met Gly Phe Phe Val Val Lys Asn Val Asn
                510                 515                 520
```

```
gct ttg gcc tgc cgc tac cga taa gctatcctca cact                    1685
Ala Leu Ala Cys Arg Tyr Arg
            525
```

<210> 5
<211> 528
<212> PRT
<213> Homo sapiens

<400> 5

```
Met Pro Ser Ser Asn Ser Arg Pro Pro Ala Cys Leu Ala Pro Gly Ala
  1               5                  10                  15
Leu Tyr Leu Ala Leu Leu Leu His Leu Ser Leu Ser Ser Gln Ala Gly
             20                  25                  30
Asp Arg Arg Pro Leu Pro Val Asp Arg Ala Ala Gly Leu Lys Glu Lys
         35                  40                  45
Thr Leu Ile Leu Leu Asp Val Ser Thr Lys Asn Pro Val Arg Thr Val
     50                  55                  60
Asn Glu Asn Phe Leu Ser Leu Gln Leu Asp Pro Ser Ile Ile His Asp
 65                  70                  75                  80
Gly Trp Leu Asp Phe Leu Ser Ser Lys Arg Leu Val Thr Leu Ala Arg
                 85                  90                  95
Gly Leu Ser Pro Ala Phe Leu Arg Phe Gly Gly Lys Arg Ala Asp Phe
            100                 105                 110
Leu Gln Phe Gln Asn Leu Arg Asn Pro Ala Lys Ser Arg Gly Gly Pro
        115                 120                 125
Gly Pro Asp Tyr Tyr Leu Lys Asn Tyr Glu Asp Ala Arg Ser Leu Asp
    130                 135                 140
Lys Leu Tyr Asn Phe Ala Asp Cys Ser Gly Leu His Leu Ile Phe Ala
145                 150                 155                 160
Leu Asn Ala Leu Arg Arg Asn Pro Asn Asn Ser Trp Asn Ser Ser Ser
                165                 170                 175
Ala Leu Ser Leu Leu Lys Tyr Ser Ala Ser Lys Lys Tyr Asn Ile Ser
            180                 185                 190
Trp Glu Leu Gly Asn Glu Pro Asn Asn Tyr Arg Thr Met His Gly Arg
        195                 200                 205
Ala Val Asn Gly Ser Gln Leu Gly Lys Asp Tyr Ile Gln Leu Lys Ser
    210                 215                 220
Leu Leu Gln Pro Ile Arg Ile Tyr Ser Arg Ala Ser Leu Tyr Gly Pro
225                 230                 235                 240
Asn Ile Gly Arg Pro Arg Lys Asn Val Ile Ala Leu Leu Asp Gly Phe
                245                 250                 255
Met Lys Val Ala Gly Ser Thr Val Asp Ala Val Thr Trp Gln His Cys
            260                 265                 270
Tyr Ile Asp Gly Arg Val Val Lys Val Met Asp Phe Leu Lys Thr Arg
        275                 280                 285
Leu Leu Asp Thr Leu Ser Asp Gln Ile Arg Lys Ile Gln Lys Val Val
```

```
    290                 295                 300
Asn Thr Tyr Thr Pro Gly Lys Lys Ile Trp Leu Glu Gly Val Val Thr
305                 310                 315                 320
Thr Ser Ala Gly Gly Thr Asn Asn Leu Ser Asp Ser Tyr Ala Ala Gly
                325                 330                 335
Phe Leu Trp Leu Asn Thr Leu Gly Met Leu Ala Asn Gln Gly Ile Asp
            340                 345                 350
Val Val Ile Arg His Ser Phe Phe Asp His Gly Tyr Asn His Leu Val
        355                 360                 365
Asp Gln Asn Phe Asn Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Tyr
    370                 375                 380
Lys Arg Leu Ile Gly Pro Lys Val Leu Ala Val His Val Ala Gly Leu
385                 390                 395                 400
Gln Arg Lys Pro Arg Pro Gly Arg Val Ile Arg Asp Lys Leu Arg Ile
                405                 410                 415
Tyr Ala His Cys Thr Asn His His Asn His Asn Tyr Val Arg Gly Ser
            420                 425                 430
Ile Thr Leu Phe Ile Ile Asn Leu His Arg Ser Arg Lys Lys Ile Lys
        435                 440                 445
Leu Ala Gly Thr Leu Arg Asp Lys Leu Val His Gln Tyr Leu Leu Gln
    450                 455                 460
Pro Tyr Gly Gln Glu Gly Leu Lys Ser Lys Ser Val Gln Leu Asn Gly
465                 470                 475                 480
Gln Pro Leu Val Met Val Asp Asp Gly Thr Leu Pro Glu Leu Lys Pro
                485                 490                 495
Arg Pro Leu Arg Ala Gly Arg Thr Leu Val Ile Pro Pro Val Thr Met
            500                 505                 510
Gly Phe Phe Val Val Lys Asn Val Asn Ala Leu Ala Cys Arg Tyr Arg
        515                 520                 525
```

<210> 6
<211> 1511
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (85)..(1497)

<400> 6
aattagtcac tagcgagacc cagtaggaag agaggtttaa atcagaggga ttgaatgagg 60

gtgctttgtg ccttcccctga agcc atg ccc tcc agc aac tcc cgc ccc ccc 111
Met Pro Ser Ser Asn Ser Arg Pro Pro
1 5

```
gcg tgc cta gcc ccg ggg gct ctc tac ttg gct ctg ttg ctc cat ctc   159
Ala Cys Leu Ala Pro Gly Ala Leu Tyr Leu Ala Leu Leu Leu His Leu
 10                  15                  20                  25

tcc ctt tcc tcc cag gct gga gac agg aga ccc ttg cct gta gac aga   207
Ser Leu Ser Ser Gln Ala Gly Asp Arg Arg Pro Leu Pro Val Asp Arg
                 30                  35                  40

gct gca ggt ttg aag gaa aag acc ctg att cta ctt gat gtg agc acc   255
Ala Ala Gly Leu Lys Glu Lys Thr Leu Ile Leu Leu Asp Val Ser Thr
             45                  50                  55

aag aac cca gtc agg aca gtc aat gag aac ttc ctc tct ctg cag ctg   303
Lys Asn Pro Val Arg Thr Val Asn Glu Asn Phe Leu Ser Leu Gln Leu
         60                  65                  70

gat ccg tcc atc att cat gat ggc tgg ctc gat ttc cta agc tcc aag   351
Asp Pro Ser Ile Ile His Asp Gly Trp Leu Asp Phe Leu Ser Ser Lys
     75                  80                  85

cgc ttg gtg acc ctg gcc cgg gga ctt tcg ccc gcc ttt ctg cgc ttc   399
Arg Leu Val Thr Leu Ala Arg Gly Leu Ser Pro Ala Phe Leu Arg Phe
 90                  95                  100                 105

ggg ggc aaa agg gcc gac ttc ctg cag ttc cag aac ctg agg aac ccg   447
Gly Gly Lys Arg Ala Asp Phe Leu Gln Phe Gln Asn Leu Arg Asn Pro
                110                 115                 120

gcg aaa agc cgc ggg ggc ccg ggc ccg gat tac tat ctc aaa aac tat   495
Ala Lys Ser Arg Gly Gly Pro Gly Pro Asp Tyr Tyr Leu Lys Asn Tyr
            125                 130                 135

gag gat gag cca aat aac tat cgg acc atg cat ggc cgg gca gta aat   543
Glu Asp Glu Pro Asn Asn Tyr Arg Thr Met His Gly Arg Ala Val Asn
        140                 145                 150

ggc agc cag ttg gga aag gat tac atc cag ctg aag agc ctg ttg cag   591
Gly Ser Gln Leu Gly Lys Asp Tyr Ile Gln Leu Lys Ser Leu Leu Gln
    155                 160                 165

ccc atc cgg att tat tcc aga gcc agc tta tat ggc cct aat att ggg   639
Pro Ile Arg Ile Tyr Ser Arg Ala Ser Leu Tyr Gly Pro Asn Ile Gly
170                 175                 180                 185

cgg ccg agg aag aat gtc atc gcc ctc cta gat gga ttc atg aag gtg   687
Arg Pro Arg Lys Asn Val Ile Ala Leu Leu Asp Gly Phe Met Lys Val
                190                 195                 200
```

```
gca gga agt aca gta gat gca gtt acc tgg caa cay tgc tac att gat   735
Ala Gly Ser Thr Val Asp Ala Val Thr Trp Gln His Cys Tyr Ile Asp
            205                 210                 215

ggc cgg gtg gtc aag gtg atg gac ttc ctg aaa act cgc ctg tta gac   783
Gly Arg Val Val Lys Val Met Asp Phe Leu Lys Thr Arg Leu Leu Asp
        220                 225                 230

aca ctc tct gac cag att agg aaa att cag aaa gtg gtt aat aca tac   831
Thr Leu Ser Asp Gln Ile Arg Lys Ile Gln Lys Val Val Asn Thr Tyr
    235                 240                 245

act cca gga aag aag att tgg ctt gaa ggt gtg gtg acc acc tca gct   879
Thr Pro Gly Lys Lys Ile Trp Leu Glu Gly Val Val Thr Thr Ser Ala
250                 255                 260                 265

gga ggc aca aac aat cta tcc gat tcc tat gct gca gga ttc tta tgg   927
Gly Gly Thr Asn Asn Leu Ser Asp Ser Tyr Ala Ala Gly Phe Leu Trp
                270                 275                 280

ttg aac act tta gga atg ctg gcc aat cag ggc att gat gtc gtg ata   975
Leu Asn Thr Leu Gly Met Leu Ala Asn Gln Gly Ile Asp Val Val Ile
            285                 290                 295

cgg cac tca ttt ttt gac cat gga tac aat cac ctc gtg gac cag aat  1023
Arg His Ser Phe Phe Asp His Gly Tyr Asn His Leu Val Asp Gln Asn
        300                 305                 310

ttt aac cca tta cca gac tac tgg ctc tct ctc ctc tac aag cgc ctg  1071
Phe Asn Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Tyr Lys Arg Leu
    315                 320                 325

atc ggc ccc aaa gtc ttg gct gtg cat gtg gct ggg ctc cag cgg aag  1119
Ile Gly Pro Lys Val Leu Ala Val His Val Ala Gly Leu Gln Arg Lys
330                 335                 340                 345

cca cgg cct ggc cga gtg atc cgg gac aaa cta agg att tat gct cac  1167
Pro Arg Pro Gly Arg Val Ile Arg Asp Lys Leu Arg Ile Tyr Ala His
                350                 355                 360

tgc aca aac cac cac aac cac aac tac gtt cgt ggg tcc att aca ctt  1215
Cys Thr Asn His His Asn His Asn Tyr Val Arg Gly Ser Ile Thr Leu
            365                 370                 375

ttt atc atc aac ttg cat cga tca aga aag aaa atc aag ctg gct ggg  1263
Phe Ile Ile Asn Leu His Arg Ser Arg Lys Lys Ile Lys Leu Ala Gly
        380                 385                 390
```

```
act ctc aga gac aag ctg gtt cac cag tac ctg ctg cag ccc tat ggg   1311
Thr Leu Arg Asp Lys Leu Val His Gln Tyr Leu Leu Gln Pro Tyr Gly
    395                 400                 405

cag gag ggc cta aag tcc aag tca gtg caa ctg aat ggc cag ccc tta   1359
Gln Glu Gly Leu Lys Ser Lys Ser Val Gln Leu Asn Gly Gln Pro Leu
410                 415                 420                 425

gtg atg gtg gac gac ggg acc ctc cca gaa ttg aag ccc cgc ccc ctt   1407
Val Met Val Asp Asp Gly Thr Leu Pro Glu Leu Lys Pro Arg Pro Leu
                430                 435                 440

cgg gcc ggc cgg aca ttg gtc atc cct cca gtc acc atg ggc ttt ttt   1455
Arg Ala Gly Arg Thr Leu Val Ile Pro Pro Val Thr Met Gly Phe Phe
            445                 450                 455

gtg gtc aag aat gtc aat gct ttg gcc tgc cgc tac cga taa           1497
Val Val Lys Asn Val Asn Ala Leu Ala Cys Arg Tyr Arg
        460                 465                 470

gctatcctca cact                                                  1511
```

<210> 7
<211> 470
<212> PRT
<213> Homo sapiens

<400> 7

```
Met Pro Ser Ser Asn Ser Arg Pro Pro Ala Cys Leu Ala Pro Gly Ala
  1               5                  10                  15
Leu Tyr Leu Ala Leu Leu Leu His Leu Ser Leu Ser Ser Gln Ala Gly
             20                  25                  30
Asp Arg Arg Pro Leu Pro Val Asp Arg Ala Ala Gly Leu Lys Glu Lys
         35                  40                  45
Thr Leu Ile Leu Leu Asp Val Ser Thr Lys Asn Pro Val Arg Thr Val
     50                  55                  60
Asn Glu Asn Phe Leu Ser Leu Gln Leu Asp Pro Ser Ile Ile His Asp
 65                  70                  75                  80
Gly Trp Leu Asp Phe Leu Ser Ser Lys Arg Leu Val Thr Leu Ala Arg
                 85                  90                  95
Gly Leu Ser Pro Ala Phe Leu Arg Phe Gly Gly Lys Arg Ala Asp Phe
            100                 105                 110
Leu Gln Phe Gln Asn Leu Arg Asn Pro Ala Lys Ser Arg Gly Gly Pro
        115                 120                 125
Gly Pro Asp Tyr Tyr Leu Lys Asn Tyr Glu Asp Glu Pro Asn Asn Tyr
    130                 135                 140
Arg Thr Met His Gly Arg Ala Val Asn Gly Ser Gln Leu Gly Lys Asp
```

145 150 155 160
Tyr Ile Gln Leu Lys Ser Leu Leu Gln Pro Ile Arg Ile Tyr Ser Arg
165 170 175
Ala Ser Leu Tyr Gly Pro Asn Ile Gly Arg Pro Arg Lys Asn Val Ile
180 185 190
Ala Leu Leu Asp Gly Phe Met Lys Val Ala Gly Ser Thr Val Asp Ala
195 200 205
Val Thr Trp Gln His Cys Tyr Ile Asp Gly Arg Val Val Lys Val Met
210 215 220
Asp Phe Leu Lys Thr Arg Leu Leu Asp Thr Leu Ser Asp Gln Ile Arg
225 230 235 240
Lys Ile Gln Lys Val Val Asn Thr Tyr Thr Pro Gly Lys Lys Ile Trp
245 250 255
Leu Glu Gly Val Val Thr Thr Ser Ala Gly Gly Thr Asn Asn Leu Ser
260 265 270
Asp Ser Tyr Ala Ala Gly Phe Leu Trp Leu Asn Thr Leu Gly Met Leu
275 280 285
Ala Asn Gln Gly Ile Asp Val Val Ile Arg His Ser Phe Phe Asp His
290 295 300
Gly Tyr Asn His Leu Val Asp Gln Asn Phe Asn Pro Leu Pro Asp Tyr
305 310 315 320
Trp Leu Ser Leu Leu Tyr Lys Arg Leu Ile Gly Pro Lys Val Leu Ala
325 330 335
Val His Val Ala Gly Leu Gln Arg Lys Pro Arg Pro Gly Arg Val Ile
340 345 350
Arg Asp Lys Leu Arg Ile Tyr Ala His Cys Thr Asn His His Asn His
355 360 365
Asn Tyr Val Arg Gly Ser Ile Thr Leu Phe Ile Ile Asn Leu His Arg
370 375 380
Ser Arg Lys Lys Ile Lys Leu Ala Gly Thr Leu Arg Asp Lys Leu Val
385 390 395 400
His Gln Tyr Leu Leu Gln Pro Tyr Gly Gln Glu Gly Leu Lys Ser Lys
405 410 415
Ser Val Gln Leu Asn Gly Gln Pro Leu Val Met Val Asp Asp Gly Thr
420 425 430
Leu Pro Glu Leu Lys Pro Arg Pro Leu Arg Ala Gly Arg Thr Leu Val
435 440 445
Ile Pro Pro Val Thr Met Gly Phe Phe Val Val Lys Asn Val Asn Ala
450 455 460
Leu Ala Cys Arg Tyr Arg
465 470

<210> 8
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PRIMER pr22-04

<400> 8
aggccaaagc attgacattc ttgaccacaa 30

<210> 9
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR22-10

<400> 9
aatggccagc ccttagtgat ggtgg 25

<210> 10
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR22-11

<400> 10
tcggtagcgg caggccaaag cattg 25

<210> 11
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR22-13

<400> 11
tcccgtcgtc caccatcact aaggg 25

<210> 12
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR22-90

<400> 12
cttacagatt catgaaggtg gcagg 25

<210> 13
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR22-91

<400> 13
ctgattggcc agcattccta aagtg 25

<210> 14
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR23-24

<400> 14
tgtagcaatg ttgccaggta agtgc 25

<210> 15
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR23-26

<400> 15
acttcctgcc accttcatga atctg 25

<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR23-41

<400> 16
tggctcgatt tcctaagctc caagc 25

<210> 17
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR23-48

<400> 17
gcaacaggct cttcagctgg atgta 25

<210> 18
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer PR23-50

<400> 18
atgtaatcct ttcccaactg gctgc 25

<210> 19
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer AP1

<400> 19
ccatcctaat acgactcact atagggc 27

<210> 20
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer AP2

<400> 20
actcactata gggctcgagc ggc 23

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/004645 A3

(51) International Patent Classification[7]: C12N 15/56, 9/24, C12Q 1/68, A61K 38/47, 48/00, G01N 33/50

(21) International Application Number: PCT/EP01/08094

(22) International Filing Date: 12 July 2001 (12.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00202442.0 12 July 2000 (12.07.2000) EP

(71) Applicant *(for all designated States except US)*: **VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW** [BE/BE]; Rijvisschestraat 120, B-9052 Zwijnaarde (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DAVID, Guido** [BE/BE]; Koningin Elisabethlaan 11, B-Belgium Leuven (BE). **DÜRR, Joachim** [DE/BE]; Bergstraat 19, B-1970 Wezembeek-Oppem (BE).

(74) Common Representative: **VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW**; Rijvisschestraat 120, B-9052 Zwijnaarde (BE).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

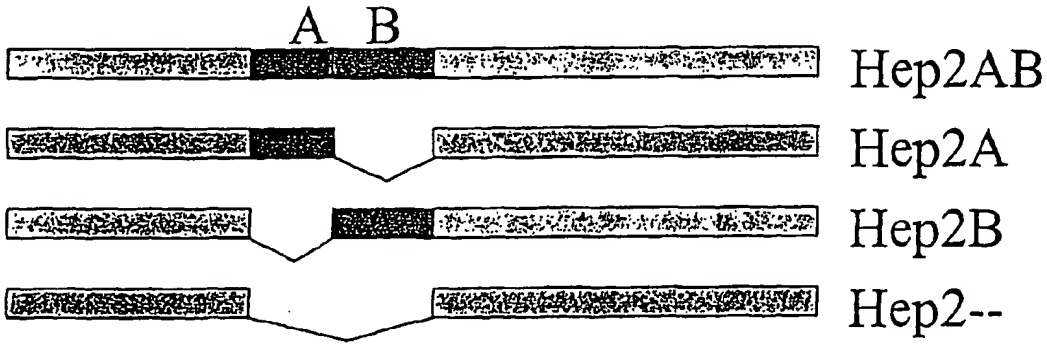
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
17 October 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: A SECOND HUMAN HEPARANASE, AND SPLICE VARIANTS THEREOF, WITH A PREDOMINANT EXPRESSION IN SKELETAL MUSCLE, HEART AND PANCREAS



(57) Abstract: The present invention relates to the field of carbohydrates and more specifically to the field of heparan sulphate proteoglycans. A polynucleotide has been identified encoding heparanase activity. This is the second gene encoding heparanase catalytic activity which is identified in humans. Several splice variants of said gene have been identified with a specific expression pattern in skeletal muscle, heart and pancreas.

WO 02/004645 A3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 01/08094

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/56 C12N9/24 C12Q1/68 A61K38/47 A61K48/00
G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, SEQUENCE SEARCH, EMBL, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! EMBL; 28 October 1998 (1998-10-28) STRAUSBERG R: "qg97h02.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1843155 3', mRNA sequence" Database accession no. AI222323 XP002155088 cited in the application	2,4,5
Y	Note: 100.0% nt seq identity with SEQ ID NO:1 in 135 nt overlap (244-378:1847-1713). the whole document	1,3,6-9, 12-15
Y	WO 99 43830 A (FAIRBANKS MICHAEL B ;HEINRIKSON ROBERT L (US); MILDNER ANA M (US);) 2 September 1999 (1999-09-02) the whole document page 9, line 13-18	1,3,6-9, 12-15
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 July 2002	19/07/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	van de Kamp, M

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KOSIR MA ET AL.: "Human prostate carcinoma cells produce extracellular heparanase" JOURNAL OF SURGICAL RESEARCH, vol. 67, no. 1, January 1997 (1997-01), pages 98-105, XP002155605 abstract page 102, right-hand column, line 23 -page 103, left-hand column, line 15 ---	1,3,8,9, 13,14
X	KOSIR M A ET AL.: "Degradation of basement membrane by prostate tumor heparanase" JOURNAL OF SURGICAL RESEARCH, vol. 81, no. 1, January 1999 (1999-01), pages 42-47, XP002155606 abstract page 45, right-hand column, line 13 -page 46, left-hand column, line 18 ---	1,3,8,9, 13,14
X	DATABASE EMBL 'Online! EMBL; 18 October 1999 (1999-10-18) HATTORI M ET AL.: "Homo sapiens genomic DNA, 21q region, clone: B2289H10 A012(-21)" Database accession no. AG019564 XP002155089 Note: 98.9% nt seq identity with SEQ ID NO:1 in 182 nt overlap (149-330:1666-1847). the whole document ---	2,4
A	WO 00 03036 A (AYAL HERSHKOVITZ MATY ;HADASIT MED RES SERVICE (IL); PECKER IRIS () 20 January 2000 (2000-01-20) cited in the application the whole document ---	
P,X	WO 01 00643 A (ITZHAKI HANAN ;MICHAL ISRAEL (IL); PECKER IRIS (IL); INSIGHT STRAT) 4 January 2001 (2001-01-04) Note: 99.7% nt seq identity of SEQ ID NO:1 with SEQ ID NO:1 in 1804 nt overlap (15-1818:44-1847), 99.5% aa seq identity of SEQ ID NO:3 with SEQ ID NO:2 in 582 aa overlap (11-592:1-582). the whole document --- -/--	1-9, 12-15

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 01 46392 A (MCKENZIE EDWARD ALEXANDER ;OXFORD GLYCOSCIENCES UK LTD (GB); STAMP) 28 June 2001 (2001-06-28) Note: 99.6% nt seq identity of SEQ ID NO:1 with SEQ ID NO:1 in 1847 nt overlap (547-2393:1-1847), 99.3% aa seq identity of SEQ ID NO:2 with SEQ ID NO:2 in 582 aa overlap (11-592:1-582). the whole document	1-9, 12-15
P,X	WO 01 21814 A (MERCK PATENT GMBH ;DUECKER KLAUS (DE); SIRRENBERG CHRISTIAN (DE)) 29 March 2001 (2001-03-29) Note: 99.6% nt seq identity of SEQ ID NO:1 with SEQ ID NO:1 in 1779 nt overlap (1-1779:55-1833), 99.1% aa seq identity of SEQ ID NO:2 with SEQ ID NO:2 in 582 aa overlap (11-592:1-582). the whole document	1-9, 12-15
P,X	WO 01 48161 A (SCHERING AG ;WEISS BERTRAM (DE); SIEMEISTER GERHARD (DE)) 5 July 2001 (2001-07-05) Note: 99.5% nt seq identity of SEQ ID NO:1 with SEQ ID NO:1 in 1010 nt overlap (484-1493:838-1847), 88.6% aa seq identity of SEQ ID NO:2 with SEQ ID NO:2 in 550 aa overlap (1-492:33-582). the whole document	1-9, 12-15
E	WO 01 79253 A (HUMAN GENOME SCIENCES INC ;EBNER REINHARD (US); RUBEN STEVEN M (US) 25 October 2001 (2001-10-25) Note: 99.7% nt seq identity of SEQ ID NO:15 with SEQ ID NO:1 in 1779 nt overlap (1-1779:55-1833), 99.5% aa seq identity of SEQ ID NO:13 with SEQ ID NO:2 in 582 aa overlap (11-592:1-582). the whole document page 13, line 14 -page 25, line 9	1-9, 12-15
E	WO 01 77341 A (SMETS GERDA MARTHA CORNELIA ;JANSSEN PHARMACEUTICA NV (BE); SPRENG) 18 October 2001 (2001-10-18) Note: 99.6% nt seq identity with SEQ ID NO:1 in 1485 nt overlap (18-1502:363-1847), 99.7% aa seq identity with SEQ ID NO:2 in 331 aa overlap (1-331:252-582). the whole document	1-9, 12-15

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 01 81569 A (HEINRIKSON ROBERT L ;UPJOHN CO (US); BIENKOWSKI MICHAEL J (US)) 1 November 2001 (2001-11-01) Note: 99.9% nt seq identity of SEQ ID NO:1 with SEQ ID NO:1 in 1010 nt overlap (634-1643:838-1847), 89.5% aa seq identity of SEQ ID NO:2 with SEQ ID NO:2 in 582 aa overlap (1-582:11-534). the whole document	1-9, 12-15
T	MCKENZIE E ET AL.: "Cloning and expression profiling of hpa2, a novel mammalian heparanase family member" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 276, no. 3, 5 October 2000 (2000-10-05), pages 1170-1177, XP002155087 Note: 99.7 % nt seq ident with SEQ ID NO:1 in 1847 nt overlap, 99.5 % aa seq ident with SEQ ID NO:2. the whole document	1-9, 12-15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 01/08094

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 12 and 13 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: -
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

Continuation of Box I.2

Present claims 10 and 11 relate to a molecule defined by reference to a desirable characteristic or property, namely that it is and antagonist or an agonist of the heparanase according to claim 1. The claims cover all molecules having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such molecules. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the molecules by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to molecules that are anti-heparanase antibodies or antisense molecules.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 01/08094

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9943830 A	02-09-1999	AU 2759199 A	15-09-1999
		EP 1060252 A2	20-12-2000
		JP 2002504376 T	12-02-2002
		WO 9943830 A2	02-09-1999
		US 6387643 B1	14-05-2002
WO 0003036 A	20-01-2000	US 6190875 B1	20-02-2001
		AU 4869799 A	01-02-2000
		EP 1097241 A1	09-05-2001
		NO 20010136 A	09-03-2001
		WO 0003036 A1	20-01-2000
WO 0100643 A	04-01-2001	AU 5244800 A	31-01-2001
		EP 1212341 A1	12-06-2002
		WO 0100643 A2	04-01-2001
		NO 20015526 A	18-12-2001
WO 0146392 A	28-06-2001	AU 2207801 A	03-07-2001
		WO 0146392 A2	28-06-2001
WO 0121814 A	29-03-2001	WO 0121814 A1	29-03-2001
		EP 1214423 A1	19-06-2002
WO 0148161 A	05-07-2001	AU 3013901 A	09-07-2001
		WO 0148161 A2	05-07-2001
WO 0179253 A	25-10-2001	AU 3645901 A	07-08-2001
		AU 3646001 A	07-08-2001
		AU 4140501 A	07-08-2001
		AU 4140601 A	07-08-2001
		AU 4140701 A	07-08-2001
		AU 4140801 A	07-08-2001
		AU 4140901 A	07-08-2001
		AU 4141001 A	07-08-2001
		AU 4141101 A	20-08-2001
		AU 4141201 A	07-08-2001
		AU 4141301 A	07-08-2001
		AU 4141401 A	07-08-2001
		AU 4141501 A	07-08-2001
		AU 4141601 A	07-08-2001
		AU 4141701 A	07-08-2001
		AU 4141801 A	07-08-2001
		AU 4313701 A	14-08-2001
		AU 4719001 A	07-08-2001
		AU 4905201 A	07-08-2001
		AU 4905301 A	07-08-2001
		AU 5076701 A	07-08-2001
		AU 5076801 A	07-08-2001
		AU 5076901 A	07-08-2001
		AU 5077001 A	07-08-2001
		AU 5077101 A	20-08-2001
		AU 5077201 A	07-08-2001
		AU 5287801 A	07-08-2001
		AU 5287901 A	07-08-2001
		AU 5332301 A	30-10-2001
		AU 5516201 A	07-08-2001
		AU 5791201 A	07-08-2001
		AU 6096901 A	07-08-2001

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 01/08094

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0179253 A		AU 6289901 A	18-10-2001
		WO 0155300 A2	02-08-2001
		WO 0155301 A2	02-08-2001
		WO 0155302 A2	02-08-2001
		WO 0155303 A2	02-08-2001
		WO 0155304 A2	02-08-2001
		WO 0155305 A2	02-08-2001
		WO 0155306 A2	02-08-2001
		WO 0155162 A1	02-08-2001
		WO 0155307 A2	02-08-2001
		WO 0154472 A2	02-08-2001
		WO 0155364 A2	02-08-2001
		WO 0155308 A2	02-08-2001
		WO 0155387 A1	02-08-2001
		WO 0155309 A2	02-08-2001
		WO 0154733 A1	02-08-2001
		WO 0155200 A1	02-08-2001
		WO 0155310 A2	02-08-2001
WO 0177341 A	18-10-2001	AU 5478001 A	23-10-2001
		WO 0177341 A2	18-10-2001
WO 0181569 A	01-11-2001	AU 5127801 A	07-11-2001
		WO 0181569 A2	01-11-2001
		US 2002064853 A1	30-05-2002